(* TWO-WAY *)

17

Elementary Analysis is often a Beginning

Introduction 17-1

17A. Elementary analysis and associated plots 17-6
- associated plots 17-9

17B. Using natural order to look further 17-16
- we need not stop 17-19

17C. Looking under the time trends 17-23

17D. Looking at the time trends 17-31
- can we see the whole fit? 17-36

17E. Bent plots 17-38
bent plot, at 17-43

17F. Another step 17-45

17G. Our analysis summarized 17-53

****

17H. How far have we come? 17-62

****

17P. Problems for chapter 17 17-64

Exh.	Page
1	17-3
2	17-4
3	17-5
A	
4	17-7
5	17-8
6	17-11
7	17-12
8	17-13
9	17-14
B	
10	17-18
11	17-21
12	17-22
C	
13	17-24
14	17-25
15	17-27
16	17-28
17	17-29/30
D	
18	17-33
19	17-34
20	17-35
21	17-37
E	
22	17-39
23	17-40
24	17-42
25	17-44
F	
26	17-46
27	17-48
28	17-49
29	17-51
30	17-52
G	
31	17-55
32	17-56
33	17-58
34	17-59
35	17-60
H	
P	

17

Elementary analysis is often a beginning

It is time now to turn to a two-way table of moderately substantial size -- one where we are not at all sure in advance of what will happen -- and to treat elementary analysis as one new tool added to our kit. We will learn that -- in this example, as in many -- elementary analysis is an essential first step, but is far from completing the analysis.

As elsewhere, this chapter will stress our dual responsibility in response to any successful fit:

- to calculate the residuals and examine them.
- to find a way to look at our fit in as direct visual terms as possible.

Now that we have come to two-way tables, neither of these responsibilities is as easy to meet as it was for data of even simpler structure. We can work hard, however. If we do, we can be surprisingly successful.

The data will be the % Democratic vote for President -- more precisely the Democratic vote as a % of the major party vote -- in 10 adjacent southcentral Nebraska counties in the 12 presidential elections from 1920 to 1964 (data from Scammon 1965).

How shall we label the versions of the two factors? Elections are carefully and compactly labeled by years. When we have more space and wish different insights, we can also use the names of the winning candidates. Counties could be labeled by their names -- if we lived in south central Nebraska this might be a good idea. Few of us, however, will know Red Willow county from Buffalo County by their names. The best we could do would be to look at a map. The best labels we can use; then, are probably map references.

Exhibit 1 relates short and long labels, Exhibit 2 shows the short labels for counties on a map of Nebraska. Exhibit 3 gives the data -- and the sums over 10 counties. The order of counties is that which -- as we see from the map -- comes closest to stretching out the counties in the order in which they lie on the map.

A moderately hard look at this exhibit shows:

- substantial and seemingly rather consistent differences from election to election
- less notable differences, also seemingly consistent from county to county (with E7 the outstandingly different county).

Is there anything else notable here? To answer this question by placing one's nose (and eyes) a few inches above exhibit 3 and moving it around, while looking and thinking hard, would not be easy. We need to do things to the numbers.

exhibit 1 of chapter 17

Short and long labels for counties and elections

Counties		Elections	
Long	Short	Short	Long
Adams	H7	'20	Harding
Buffalo	G6	'24	Coolidge
Dawson	G4	'28	Hoover
Frontier	H3	'32	Roosevelt 1
Hall	G7	'36	Roosevelt 2
Hamilton	G8	'40	Roosevelt 3
Howard	E7	'44	Roosevelt 4
Merrick	E8	'48	Truman
Polk	E9	'52	Eisenhower 1
Red Willow	I2	'56	Eisenhower 2
		'60	Kennedy
		'64	Johnson

exhibit 2 of chapter 17

The Ten Counties on the Map of Nebraska
(see exhibit 1 for names)

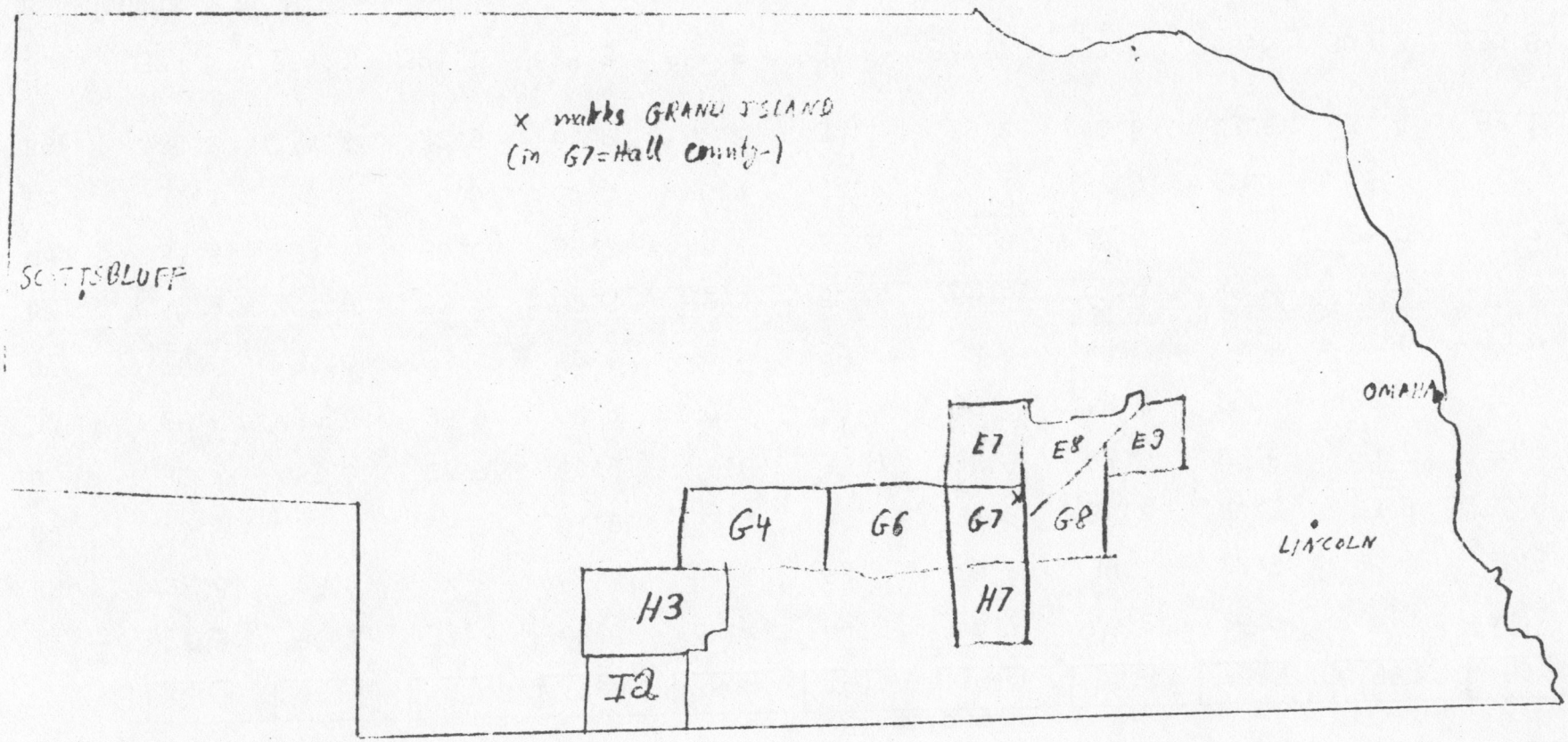


exhibit 3 of chapter 17

% of Major Party Vote for President Given to the Democratic Candidate --
10 Nebraska Counties in 12 Elections

A) DATA

	Elections											
	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64
I2	37.3	36.8	33.2	63.8	62.4	44.3	40.7	46.2	26.9	29.9	33.1	46.9
H3	27.8	28.6	27.3	61.8	54.4	33.9	29.0	39.3	22.9	28.4	26.5	43.6
G4	29.9	33.6	25.1	61.2	52.9	34.0	31.2	38.2	20.3	23.9	26.9	45.3
G6	31.3	33.0	27.3	60.9	56.6	38.9	38.8	43.8	22.8	29.7	33.9	50.1
H7	28.5	32.8	28.9	58.9	59.9	39.4	39.2	45.6	29.3	33.1	35.5	53.6
G7	26.8	31.6	33.1	62.6	55.0	38.7	38.4	44.6	30.6	33.6	36.0	55.2
E7	46.5	56.8	53.1	82.3	72.0	56.9	56.8	62.6	40.8	46.0	47.5	66.5
G8	31.5	34.5	30.6	59.7	49.1	33.6	30.3	38.5	24.0	25.0	27.3	47.3
E8	31.1	32.9	30.0	62.9	50.4	33.7	34.1	37.9	22.7	28.3	31.0	49.2
E9	34.1	34.3	32.5	64.2	52.8	38.1	39.2	40.8	23.3	31.4	33.1	51.8
Sums	324.8	354.9	321.1	638.3	565.5	391.5	377.7	437.5	263.6	309.3	330.8	509.5

B) SOURCE: R. M. Scammon 1965. America at the Polls. A handbook of American presidential election statistics 1920-1964. Governmental Affairs Institute. University of Pittsburgh Press. From pages 276-287.

17A. Elementary analysis and associated plots

The obvious first step is to make allowance for election-to-election differences. The easiest thing to do is to find the arithmetic mean -- over the 10 counties -- for each election, and then to form residuals by subtracting the corresponding election mean from each observed %. The result is shown in exhibit 4. Differences from county to county are now clearer, and it would be substantially easier to try to look beneath the surface to see what else may be going on. Again, a little arithmetic can clear the way appreciably. So let us use it.

Exhibit 5 shows the effect of finding the mean -- over 12 elections -- of the first residuals for each county, and then subtracting the corresponding such mean from each first residual. We now have a mass of second residuals that are appreciably smaller than the numbers we have been so far examining. Nothing striking leaps to the eye, though there may well be further effects to be found.

We may as well complete the elementary analysis, as we have done in exhibit 5. We shall sometimes find it convenient to call the residuals found by successive subtraction of pieces of a partial description "first residuals", "second residuals", "third residuals", and so on. We do this as a matter of convenience, not of principle -- indeed, what are "fourth residuals" by one chain of steps might be "seventh residuals" by another. We do it for one simple reason: because

exhibit 4 of chapter 17

(see exhibit 3)

First Residuals (from election means) of % Democratic

County	Election												
	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64	(Sum)
(**)	(32.5)	(35.5)	(32.1)	(63.8)	(56.6)	(39.1)	(37.8)	(43.8)	(26.4)	(30.9)	(33.1)	(51.0)	((482.6))
I2	+4.8	+1.3	+1.1	-0.0	+5.8	+5.2	+2.9	+24.	+0.5	-1.0	0.0	-4.1	(18.9)
H3	-4.7	-6.9	-4.8	-2.0	-2.2	-5.2	-8.8	-4.5	-3.5	-2.5	-6.6	-7.4	(-58.1)
G4	-2.6	-1.9	-7.0	-2.6	-3.7	-5.1	-6.6	-5.6	-6.1	-7.0	-6.2	-5.7	(-60.6)
G6	-1.2	-2.5	-4.8	-2.9	-0.0	-0.2	+1.0	-0.0	-3.6	-1.2	0.8	-0.9	(-15.5)
H7	-4.0	-2.7	-3.2	-4.9	+3.3	+0.3	+1.4	+1.8	+2.9	2.2	2.4	2.6	(+2.1)
G7	-5.7	-3.9	+1.0	-1.2	-1.6	-0.4	+0.6	+0.8	+4.2	+2.7	+2.1	+5.8	(+4.4)
E7	14.0	21.3	21.0	18.5	15.4	17.8	19.0	18.8	14.4	15.1	14.4	15.5	(205.4)
G8	-1.0	-1.0	-1.5	-4.1	-7.5	-5.5	-7.5	-5.3	-2.4	-5.9	-5.8	-3.7	(-51.0)
E8	-1.4	-2.6	-2.1	-0.9	-6.2	-5.4	-3.7	-5.9	-3.7	-3.6	-2.1	-1.8	(-39.4)
E9	+1.6	-1.3	+0.4	+0.4	-3.8	-1.0	+1.4	-3.0	-3.1	0.5	0.0	0.8	(-7.2)

Examples: 32.5 = 324.8/10, -4.0 = 28.5 - 32.5.

(Some computational errors can be found!)

(**) Election means, already subtracted.

exhibit 5 of chapter 17

Elementary Analysis, Including Second Residuals, After Subtraction of County Means

County	Election '20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64	Eff	%D
I2	3.2	-0.3	-0.5	-1.6	4.2	3.6	1.3	0.8	-1.1	-2.6	-1.6	-5.7	**1.6**	41.8
H3	0.1	-2.1	0.0	2.8	2.6	-0.4	-4.0	0.3	1.3	2.3	-1.8	-2.6	**-4.8**	35.4
G4	2.4	3.1	-2.0	2.4	1.3	-0.1	-1.6	-0.6	-1.1	-2.0	-1.2	-0.7	**-5.0**	35.2
G6	0.8	-1.2	-3.5	-1.6	1.3	1.1	2.3	1.3	-2.3	0.1	2.1	0.4	**-1.3**	38.9
H7	-4.2	-2.9	-3.4	-5.1	3.1	0.1	1.2	1.6	2.7	2.0	2.2	2.4	**0.2**	40.4
G7	-6.1	-4.3	0.6	-1.6	-2.0	-0.8	0.2	0.4	3.8	2.3	1.7	5.4	**0.4**	40.6
E7	-3.1	4.2	3.9	1.4	-1.7	0.7	1.9	1.7	-2.7	-2.0	-2.7	-1.6	**17.1**	57.3
G8	3.3	3.3	2.8	0.2	-3.2	-1.2	-3.2	-1.0	1.9	-1.6	-1.5	0.6	**-4.3**	35.9
E8	1.9	0.7	1.2	2.4	-2.9	-2.1	-0.4	-2.6	-0.4	-0.3	1.2	1.5	**-3.3**	36.9
E9	2.2	-0.7	1.0	1.0	-3.2	-0.4	2.0	-2.4	-2.5	1.1	0.6	1.4	**-0.6**	39.6
Eff	**-7.7**	**-4.7**	**-8.1**	**23.6**	**16.4**	**-1.1**	**-2.4**	**3.6**	**-13.8**	**-9.3**	**-7.1**	**10.8**	**40.2**	00.0
%D	32.5	35.5	32.1	53.8	56.6	39.1	37.8	43.8	26.4	30.9	33.1	51.0	00.0	-40.2

Examples: 40.2 = 482.6/12; +0.2 = 2.1/12; -1.3 = -15.6/12; -4.2 = -4.0 - (0.2); -1.2 = (-2.5) - (-1.3).

it gives us convenient, confusion-reducing labels for the various residuals that arise as we move along. We will say merely "residuals" whenever this seems simpler.

One natural thing to do at this stage is to look at the pattern of the values of these 120 second residuals, first as 120 numbers, and then with attention to how they are placed in the two-way table. The stem-and-leaf display of exhibit 6 shows the pattern as displayed, and goes on to find hinge and side points.

* associated plots *

Before we go on to look at the residuals in other ways, we ought to look at:

- a two-way plot of the fit. See exhibit 7 and note that the residuals are relatively larger than in our examples in the last chapter.
- a plot of residuals superimposed on that plot. See exhibit 8 for calculation of the appropriate residuals. See exhibit 9 for the plot, which shows no distinctive pattern.

Since the purpose of plotting coded residuals on top of an (invisible) two-way plot is to spot patchiness, we do not need to plot too many points. We can wisely combine groups of rows and groups of columns, plotting a single residual for all the points common to both a group of rows and a group of columns.

Rules for forming groups can be given. We shall not bother, feeling that a hard look at the two-way plot -- and common sense -- will do well enough.

When we look at exhibit 7, we are led to these groupings: Three for counties: E4 alone; I2, H7, G6, G7 and B9; E8, G8, G4 and H3. Eight for elections: '52 alone; '20, '28, '56 and '60, '24 alone, '40 and '44, '48 alone, '64 alone, '36 alone, '32 alone. Exhibit 8 applies these groupings, summarizing by median residuals.

Let us look next at the outside residuals. Three of the second residuals are beyond the side points and thus deserve special attention:

-6.1 in 1920 for G7 (this is Hall county)

-5.7 in 1964 for I2 (this is Red Willow county)

+5.4 in 1964 for G7 (Hall county, again).

exhibit 6 of chapter 17

Stem-and-leaf display of the distribution of values
of 120 second residuals from exhibit 5

(unit = 0.1%D)

to	1	-6	1
	3	-5	17
	6	-4	203
	12	-3	452 212
	28	-2	930 016 070 791 645 6
	44	-1	266 128 620 657 661 6
to	57	-0	167 048 443 743 5
120	✓ 15	+0	181 413 624 267 768
to	48	+1	263 133 379 497 922 500 143
	27	+2	702 431 448 633 842 0
	11	+3	118 339 26
	3	+4	22
to	1	+5	4
		+6	

Batch:	120 at 0.1; hinges at 30h
Hinges:	-1.6h and +1.8 midspread: 3.4h
Side points:	-5.1 and +5.2h
Outside high:	+5.4
Outside low:	-6.1 and -5.7

'exhibit 7 of chapter 17

The Fit of Exhibit 5 Plotted

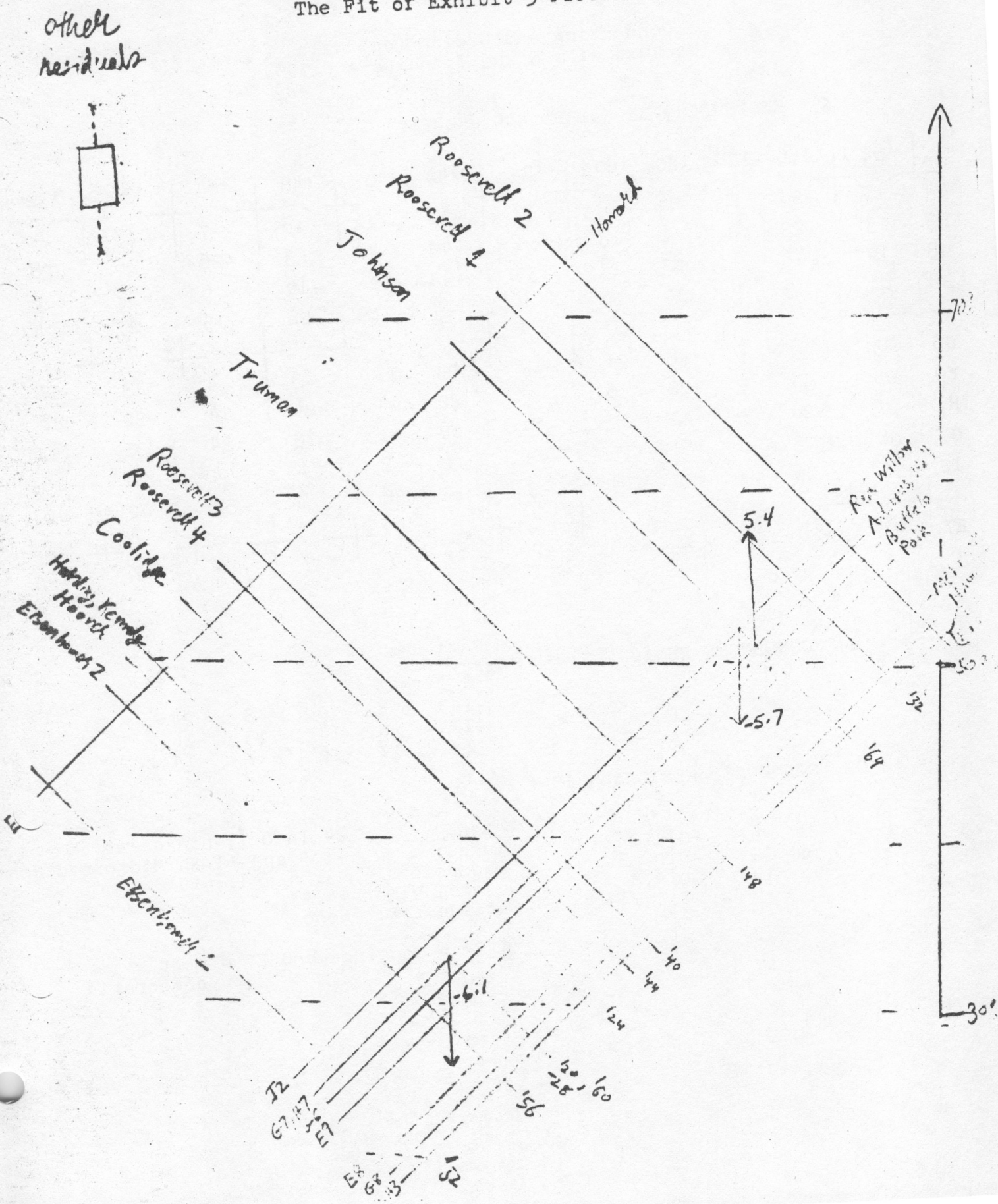


exhibit 8 of chapter 17

Condensing and Coding Residuals
(see exhibits 5 and 7, units = 0.1%D)

A) RESIDUALS, REORDERED AND GROUPED

	'52	'56	'28	'20	'60	'24	'44	'40	'48	'64	'36	'32
G4	-11	-20	-20	24	-12	31	-16	-1	-6	-7	13	24
H3	13	23	0	1	-18	-21	-40	-4	3	-26	26	28
G8	19	-16	28	33	-13	33	-32	-12	-10	6	-32	2
E8	-4	-3	12	19	12	7	-4	-21	-26	15	-29	24
G6	-23	1	-35	8	21	-12	23	11	13	4	13	-16
E9	-25	11	10	22	6	-7	20	-4	-24	14	-32	10
H7	27	20	-34	-42	22	-29	12	1	16	24	31	-51
G7	38	23	6	-61	17	-43	2	-8	4	54	-20	-16
I2	-11	-26	-5	32	-16	-3	13	36	8	-57	42	-16
E7	-27	-20	39	-31	-27	42	19	7	17	-16	-17	14

B) CONDENSED RESIDUALS -- two-way

4	0	19	-14	-8	0	-8	24
-11	7	-12	12	8	14	13	-16
-27	-24	42	13	17	-16	-17	14

C) CONDENSED RESIDUALS AS A BATCH

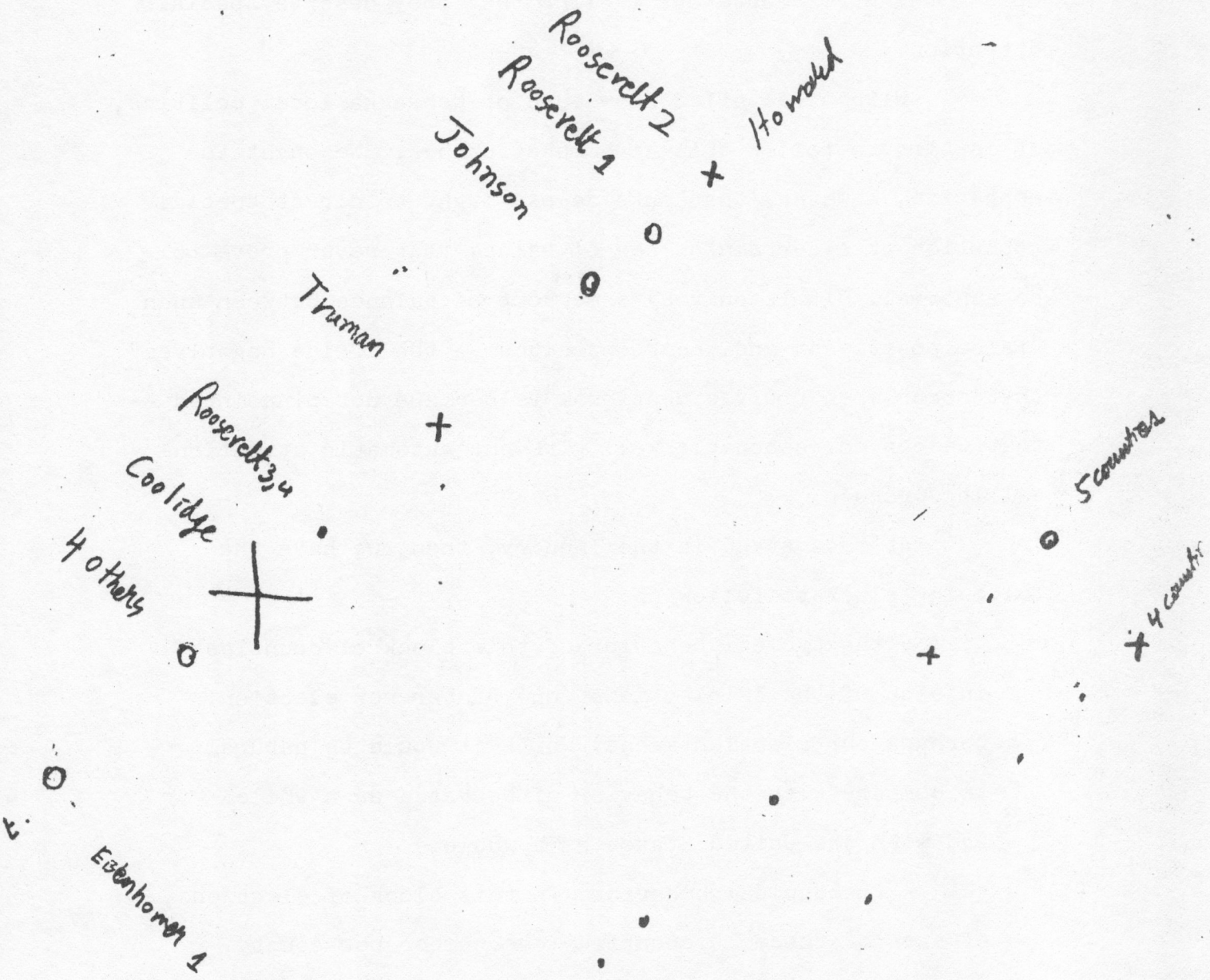
2	-2	74
8	-1	124676
✓ to 12	-0	0808
to 12	0	478
9	1	2374349
2	2	4
.	3	
1	4	2

Batch: 24 in 0.1; hinges at 6h
Hinges: -1.3 and +1.3h Midspread:26
Side points: -3.9h and +4.0
Outside high: +4.2*
Outside low: none

* '24 = Coolidge and E7 = most democratic

exhibit 9 of chapter 17

Condensed and Coded Residuals
(see exhibits 7 and 8)



This piling up of all three of these values in two counties and two elections -- the first and last of the 12 considered -- does nothing to reduce our feeling that they deserve special attention.

Without detailed knowledge of Nebraska local politics, it is hard to follow this line much further. We ought to emphasize, however, that outsideness ought to direct special attention at a certain number of values that never prove to be abnormal. It is only by some sort of balance between such "false positives" and their opposites -- the "false negatives" that occur when truly exceptional values are not pinpointed -- that we can do reasonably well with any automatic attention-calling process.

At this stage in the inquiry, then, we have the following clues to follow up:

- the typical behavior of this block of counties in each of the 12 elections (one number per election, perhaps our election means) which it would be natural to compare with the behavior of Nebraska as a whole and with the United States as a whole.

- the typical behavior, in this block of elections, of each of these 10 counties (one number per county, either county effects or count fits), which it would be natural to compare among themselves.

- the attention-deserving behavior of Hall county in 1920 and 1964, and of Red Willow county in 1964.

If our analysis were to stop here, we would certainly make these comparisons -- and then ask all sorts of questions about the differences that appeared. Why, for example, is Howard county so different (by being more than 15% more Democratic) than all the other nine counties? Does it depend less on farming? Does it do a different sort of farming? Are the national origins -- or religious associations -- of its people differently distributed? Or is there a long history of a -- once accidentally established -- strong Democratic county organization? And so on, and on, not only for this difference but for many others.

17B. Using natural order to look further

It is now time to use coded residuals in other ways. The 12 elections have a sensible order -- in time from 1920 to 1964 -- in addition to the order given to them by the corresponding effects. The 10 counties could be -- and were -- arranged in a relatively sensible order -- according to the map from I2 to E9.

Since we are so fortunate as to have these orderings, we should use them to look at the residuals. For this purpose, we need to summarize the values enough, but not too much. On one hand, the full detail of exhibit 5 is much too

detailed -- if there are patterns, they are likely to be hidden. On the other hand, the great summarization obtained by only noting the sign, + or -, of each residual -- or by using 3 classes -- is much too condensed. We need a convenient scheme of an intermediate degree of detail, such as the five-class one provided by the hinges and side points of exhibit 5. Exhibit 10 shows the results of applying such a technique with two choices: counties as columns west to east and elections as columns over time.

The double appearance of the same information, the second time with rows and columns interchanged, is important. It is easy to compare LO and HI looking down a column, but hard to do a similar thing looking along a row.

Little in this table may spring to the eye at our first glance, but if we look hard we will see quite a lot.

When we look up and down the columns of the upper table, we see a considerable tendency for nearby counties to behave somewhat similarly in each separate election; this is most striking in 1936, when the western counties tend to be high and the eastern low.

When we look up and down the columns of the lower table, we see even more. Of the four outside points -- three LO and one HI -- we find every one to be part of a clear trend with time -- toward low for I2 (Red Willow county) and toward high for H7 (Adams county) and G7 (Hall county).

exhibit 10 of chapter 17

Coded Residuals From Exhibit 5 in Two Displays

Elections as Columns

	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64
I2	·H	·	·	·	·H	·H	·	·	·	L·	·	LO·
H3	·	L·	·	·H	·H	·	L·	·	·	·H	L·	L·
G4	·H	·H	L·	·H	·	·	·	·	·	L·	·	·
G6	·	·	L·	·	·	·	·H	·	L·	·	·H	·
H7	L·	L·	L·	L·	·H	·	·	·	·H	·H	·H	·H
G7	LO·	L·	·	·	L·	·	·	·	·H	·H	·	·HI
E7	L·	·H	·H	·	L·	·	·H	·	L·	L·	L·	·
G8	·H	·H	·H	·	L·	·	L·	·	·H	·	·	·
E8	·H	·	·	·H	L·	L·	·	L·	·	·	·	·
E9	·H	·	·	·	L·	·	·H	L·	L·	·	·	·

Counties as Columns

	I2	H3	G4	G6	H7	G7	E7	G8	E8	E9
'20	·H	·	·H	·	L·	LO·	L·	·H	·H	·H
'24	·	L·	·H	·	L·	L·	·H	·H	·	·
'28	·	·	L·	L·	L·	·	·H	·H	·	·
'32	·	·H	·H	·	L·	·	·	·	·H	·
'36	·H	·H	·	·	·H	L·	L·	L·	L·	L·
'40	·H	·	·	·	·	·	·	·	L·	·
'44	·	L·	·	·H	·	·	·H	L·	L·	·H
'48	·	·	·	·	·	·	·	·	·	L·
'52	·	·	·	L·	·H	·H	L·	·H	L·	L·
'56	L·	·H	L·	·	·H	·H	L·	·	·	·
'60	·	L·	·	·H	·H	·	L·	·	·	·
'64	LO·	L·	·	·	·H	·HI	·	·	·	·

Coding used:

residual < -5.1 as LO· -5.1 ≤ residual ≤ -1.7 as L·

-1.6 ≤ residual ≤ 1.8 as · 1.8 < residual ≤ 5.2 as ·H

5.3 ≤ residual as ·HI

These trends account in a simple way for the four attention-deserving values we noted at the end of the previous section (17A) and should probably replace them in our mental summary.

If we were to stop our analysis here, we might summarize our clues as follows:

- the typical behavior of these counties over 12 elections
- the typical behavior over these elections of the 10 counties
- time trends of notable size in some counties, toward the Democrats in Adams and Hall, toward the Republicans in Red Willow
- east-west gradients in some elections.

If we were to stop here, we would make all the comparisons mentioned at the close of section 17A, ask all the questions about typical behavior that would have been asked at that point, and go on and ask many political questions about time trends.

* we need not stop *

In this example, the most prominent effects, once we clear away rows (= counties) and columns (= elections), are the apparent time trends in Adams, Hall, Red Willow, and possibly, Howard counties. We say they seemed to be there, but we have not looked AT them. Exhibit 11, which shows (second)

residuals against date for each of these four counties, remedies this failure. When we look at these four plots, it is quite clear that there were time trends. Why were they there? We ought at least to ask what we can learn easily about changes in those counties over these years.

Before we turn to ask economic, social, ethnic, or religious questions about the differences between these counties, we ought to consider grasping at any straws our source of voting data offers us. In addition to the % split of the vote, this source offers us information on total vote. While differences in party activity can affect turnout in individual elections, long-time trends in total vote probably indicate long-term changes in county populations. Accordingly it seems sensible to inquire whether differences in %D trends are associated with differences in total vote trends. Exhibit 12 gives the figures, both for individual elections and for mid-means of early and late thirds. Clearly the rates of growth of votes over some 30 years -- which presumably reflect the rates of growth of population -- are quite different for the four counties: about 32% for Adams and about 60% for Hall, as against about 17% for Red Willow and about -17% for Howard. The vote in the two counties with noticeably increasing residuals grew notably faster than the vote in the counties with noticeably decreasing residuals. Accordingly, we ought to look further.

Exhibit 11 of chapter 17

Behavior of residuals over time for 4 counties

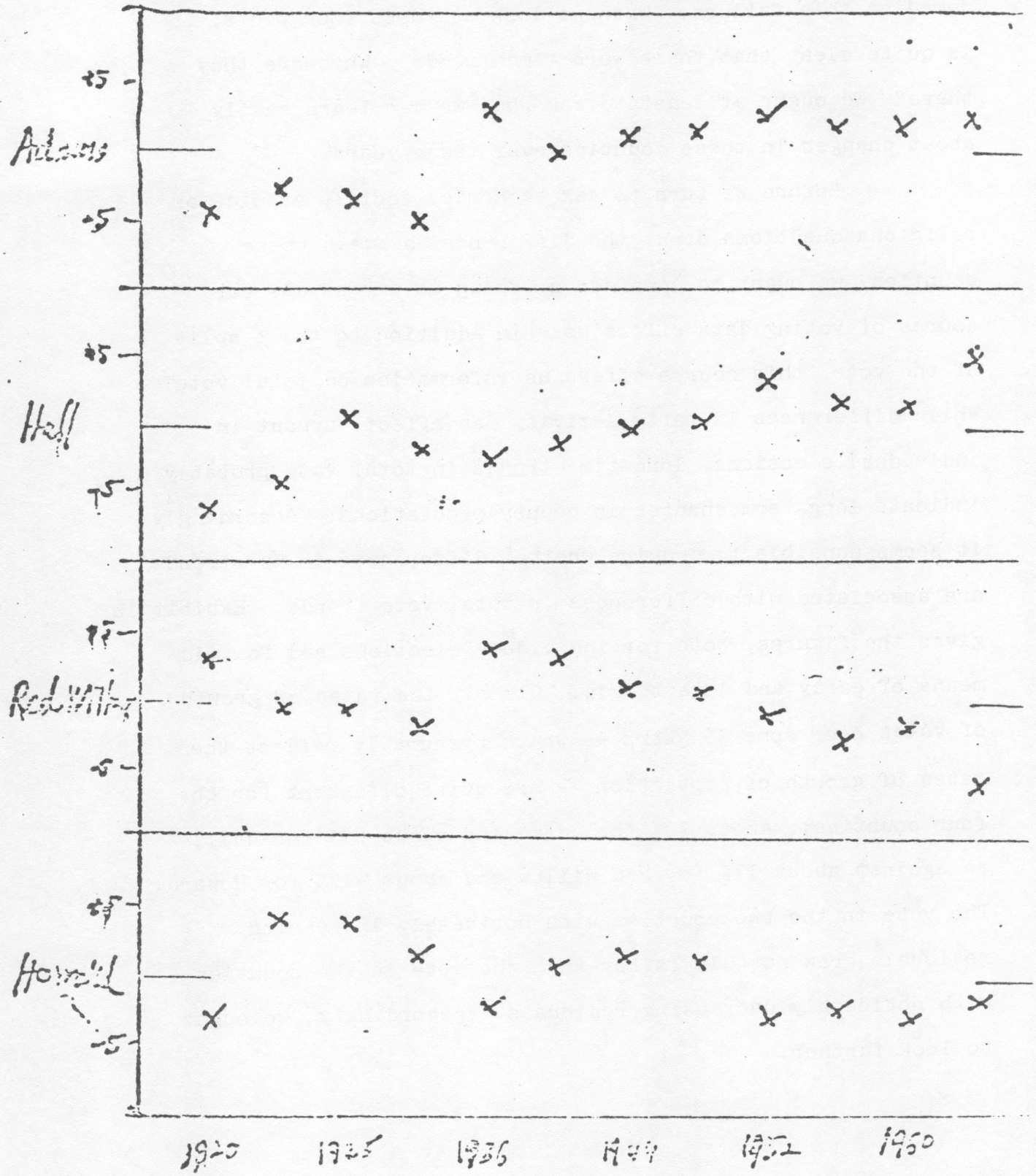


exhibit 12 of chapter 17

Total Presidential Election Vote in Four Counties

	(H7) Adams	(G7) Hall	(I2) Red Willow	(E7) Howard
1920	7,008	7,123	3,385	2,971
1924	8,516	8,525	4,329	3,455
1928	10,186	10,306	5,384	4,157
1932	9,934	10,492	5,592	4,306
1936	10,621	11,810	5,692	4,494
1940	10,941	12,099	5,600	3,934
1944	11,777	12,419	5,239	3,598
1948	10,212	10,284	4,848	3,026
1952	12,778	15,043	6,064	3,571
1956	12,234	14,351	5,431	3,150
1960	12,296	15,259	5,812	3,194
1964	12,027	14,988	5,156	3,044
(midmean 1920 to 1932)	(9,225)	(9,416)	(4,856)	(3,806)
(midmean 1952 to 1964)	(12,265)	(15,015)	(5,622)	(3,172)

Exhibit 13 presents the medians for early and late thirds of the 12 elections for both residuals and total votes. Exhibit 14 plots the change in median residual against the % change in median total vote. We see a definite appearance of increasing (= more Democratic) residuals with increasing growth in total vote. The county of Dawson, marked D to the lower right, might be enough different in behavior to deserve special consideration.

We can now add many questions about the economic, social, ethnic and religious factors which may carry the effects of different rates of population growth into different rates of trend in % Democratic.

17C. Looking under the time trends

To look still deeper, we have again to "clear the trail". We have recognized the existence of appreciable time trends; we must admit that their presence could conceal other phenomena. Are there other phenomena? To see the answer to such a question we need to strip away whatever we have recognized -- here, visible time trends. So let us begin.

To calculate how much time trend appears to be present -- if this trend follows a straight-line course -- is easy. If the trend were

-11,-9,-7,-5,-3,-1,+1,+3,+5,+7,+9,+11

which is surely straight-line and arithmetically simple, the

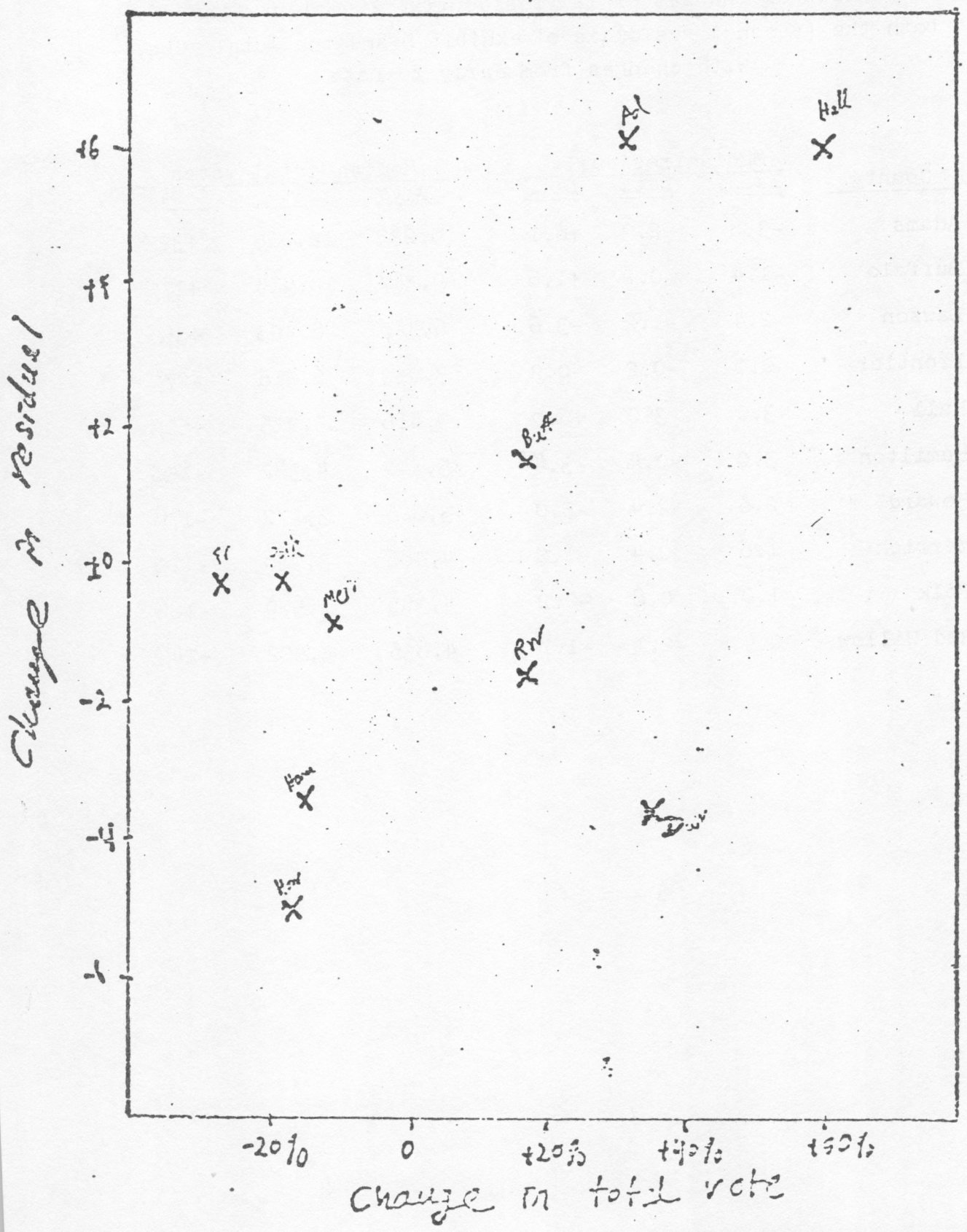
exhibit 13 of chapter 17

Medians for early (1920-1932) and late (1952-1964) thirds of the set of 12 Presidential elections for both the (second) residuals of exhibit 5 and the total vote, with changes from early to late

County	Median residuals			Median total votes		
	Early	Late	Diff	Early	Late	% Diff
H7 = Adams	-3.8	2.3	+6.1	9,225	12,265	+32%
G6 = Buffalo	-1.4	0.2	+1.6	9,347	10,914	+17%
G4 = Dawson	2.4	-1.2	-3.6	6,427	8,705	+35%
H3 = Frontier	0.0	-0.2	-0.2	3,059	2,216	-27%
G7 = Hall	-3.0	3.0	+6.0	9,416	15,015	+60%
G8 = Hamilton	3.0	-0.4	-3.4	5,140	4,382	-15%
E7 = Howard	2.6	-2.4	-5.0	3,806	3,172	-17%
E8 = Merrick	1.6	0.4	-1.2	4,486	3,980	-11%
E9 = Polk	1.0	0.8	-0.2	4,383	3,599	-18%
I2 = Red Willow	-0.4	-2.1	-1.7	4,856	5,622	+14%

exhibit 14 of chapter 17

Comparison of changes from 1920-1932 to 1952-1964
for the 10 counties



medians of the early and late thirds would be -8 and +8, respectively, so that their difference would be +16.

For Buffalo county (see exhibit 12) the observed difference of the corresponding medians, in %D, is +1.6. Clearly we may adjust Buffalo county by subtracting one-tenth of the values (-11 in 1920 to +11 in 1964) we took as a sample trend. Clearly we should act proportionately in other counties. In Howard county, for instance, we should subtract (-5.0)/16 times the simple values (that is, add 5/16 times them).

Exhibit 15 sets out the resulting adjusted (third) residuals, which appear somewhat smaller, overall, than the unadjusted (second) residuals. Indeed, as exhibit 16 shows, the midspread is reduced to about 5/7 of its earlier value, so that its square is reduced to about 1/2 what it was for second residuals. Clearly, taking out a linear time trend for each county has made the residuals quite a lot more pulled together.

Having these residuals, we should look at them. Our fit is no longer one that can be simply plotted in a two-way plot, so we shall not try to look at the residuals against such a background. Tabulated five-band residuals should meet our needs.

Exhibit 17 provides one table for each of the four natural and implied orders of the factors. Looking hard, we see:

- taking counties from Republican to Democratic:

tendency for the election of '32 to drift downward,

exhibit 15 of chapter 17

Residuals after removing linear time trends from residuals of exhibit 5
(third residuals)

County	Election												County
	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64	
I2	2.0	-1.3	-1.2	-2.1	3.9	3.5	1.4	1.1	-0.6	-1.9	-0.6	-4.5	Red Willow
H3	0.0	-2.2	-0.1	2.7	2.6	-0.4	-4.0	0.3	1.4	2.4	-1.7	-2.5	Frontier
G4	-0.1	1.1	-3.6	1.3	0.6	-0.3	-1.4	0.1	0.0	0.4	0.8	1.8	Dawson
G6	1.9	-0.3	-2.8	-1.1	1.6	1.2	2.2	1.0	-2.8	-0.6	1.2	-0.7	Buffalo
H7	0.0	0.5	-0.7	-3.2	4.2	0.5	0.8	0.5	0.8	-0.7	-1.2	-1.8	Adams
G7	-2.0	-0.9	3.2	0.3	-0.9	-0.4	-0.2	-0.7	1.9	-0.3	-1.7	1.3	Hall
E7	-6.5	1.4	1.7	-0.2	-2.6	0.4	2.2	2.6	-1.1	0.2	0.1	1.8	Howard
G8	1.0	1.4	1.3	-0.9	-3.8	-1.4	-3.0	-0.4	3.0	-0.1	0.4	2.9	Hamilton
E8	1.1	0.0	0.7	2.0	-3.1	-2.2	-0.3	-2.4	0.0	0.2	1.9	2.3	Merrick
E9	2.1	-0.8	0.9	0.9	-3.2	-0.4	2.0	-2.4	-2.4	1.2	0.7	1.5	Polk

exhibit 16 of chapter 17

Third residuals examined as a batch

A) STEM-and-LEAF -- unit 0.1%D

```
   to    1   -6|5
         .   -5|
         3   -4|05
         9   -3|628120
        20   -2|02816244845
        31   -1|32144192778
     to 55   -0|139817299434423746673167
120  ✓ 25    0|0050793965483150804228147
     to 40    1|91104437362411492298385
        17    2|107062206493
         5    3|2950
     to  1    4|2
```

Batch:	120 in 0.1; hinges at 30h
Hinges:	-1.1 and +1.3 midspread: 2.4
Sidepoints:	-3.5 and +3 7
Outside high:	+3.9, +4.2
Outside low:	-6.5, -4.5, -4.0 -3.8, -3.6

B0 COMPARISON OF 2nd AND 3rd RESIDUALS -- see exhibit 6 and A above

Residuals	Hinges*	Midspread**	$(\text{Midspread})^2$ ***
2nd	-16h and +18	3.45	11.9
3rd	-11 and +13	2.4	5.8

*unit here 0.1%D **unit here 1%D ***unit here $(1\%D)^2$

exhibit 17 of chapter 17

Coded Residuals from exhibit 15

A) COUNTIES FROM REPUBLICAN TO DEMOCRATIC -- columns are elections

Eff	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64	County
-5.0	·	·	LO·	·	·	·	L·	·	·	·	·	·H	G4
-4.8	·	L·	·	·H	·H	·	LO·	·	·H	·H	L·	L·	H3
-4.3	·	·H	·	·	LO·	L·	L·	·	·H	·	·	·H	G8
-3.3	·	·	·	·H	L·	L·	·	L·	·	·	·H	·H	E8
-1.3	·H	·	L·	·	·H	·	·H	·	L·	·	·	·	G6
-0.6	·H	·	·	·	L·	·	·H	L·	L·	·	·	·H	E9
0.2	·	·	·	L·	·H	·	·	·	·	·	L·	L·	H7
0.4	L·	·	·H	·	·	·	·	·	·H	·	L·	·H	G7
1.6	·H	L·	·	L·	·H	·HI	·H	·	·	L·	·	·	I2
17.1	LO·	·H	·H	·	L·	·	·H	·H	·	·	·	·H	E7
	-6.9	-3.9	-7.3	14.4	17.2	-0.3	-1.6	4.4	-13.0	-8.5	-6.3	11.6	

B) ELECTIONS IN CHRONOLOGICAL ORDER -- columns are counties

Elect	G4	H3	G8	E8	G6	E9	H7	G7	I2	E7	Effect
'20	·	·	·	·	·H	·H	·	L·	·H	LO·	-6.9
'24	·	L·	·H	·	·	·	·	L·	L·	·H	-3.9
'28	LO·	·	·	·	L·	·	·	·H	·	·H	-7.3
'32	·	·H	·	·H	L·	·	L·	·	L·	·	14.4
'36	·	·H	LO·	·	·H	L·	·HI	·	·HI	L·	17.2
'40	·	·	L·	L·	·	·	·	·	·H	·	-0.3
'44	L·	LO·	L·	·	·H	·H	·	·	·H	·H	-1.6
'48	·	·	·	L·	·	L·	·	·	·	·H	4.4
'52	·	·H	·H	·	L·	L·	·	·H	·	·	-13.0
'56	·	·H	·	·	·	·	·	·	L·	·	-8.5
'60	·	L·	·	·H	·	·	L·	L·	·	·	-6.3
'64	·H	L·	·H	·H	·	·H	L·	·	LO·	·H	11.6
Effect	-5.0	-4.8	-4.3	-3.3	-1.3	-0.6	0.2	0.4	1.6	17.1	

C) COUNTIES FROM WEST TO EAST -- columns are elections

County	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64	County
I2	·H	L·	·	L·	·HI	·H	·H	·	·	L·	·	LO·	Red Willow
H3	·	L·	·	·H	·H	·	LO·	·	·H	·H	L·	L·	Frontier
G4	·	·	LO·	·	·	·	L·	·	·	·	·	·H	Dawson
G6	·H	·	L·	·	·H	·	·H	·	L·	·	·	·	Buffalo
H7	·	·	·	L·	·HI	·	·	·	·	·	L·	L·	Adams
G7	L·	·	·H	·	·	·	·	·	·H	·	L·	·	Hall
E7	LO·	·H	·H	·	L·	·	·H	·H	·	·	·	·H	Howard
G8	·	·H	·	·	LO·	L·	L·	·	·H	·	·	·H	Hamilton
E8	·	·	·	·H	L·	L·	·	L·	·	·	·H	·H	Merrick
E9	·H	·	·	·	L·	·	·H	L·	L·	·	·	·H	Polk
Eff	-6.9	-3.9	-7.3	14.4	17.2	-0.3	-1.6	4.4	-13.0	-8.5	-6.3	11.6	

D) ELECTIONS FROM REPUBLICAN TO DEMOCRATIC -- columns are counties

Eff	-5.0	-4.8	-4.3	-3.3	-1.3	-0.6	0.2	0.4	1.6	17.1	Years
-13.0	·	·H	·H	·	L·	L·	·	·H	·	·	'52
-8.5	·	·H	·	·	·	·	·	·	L·	·	'56
-7.3	LO·	·	·	·	L·	·	·	·H	·	·H	'28
-6.9	·	·	·	·	·H	·H	·	L·	·H	LO·	'20
-6.3	·	L·	·	·H	·	·	L·	L·	·	·	'60
-3.9	·	L·	·H	·	·	·	·	·	L·	·H	'24
-1.6	L·	LO·	L·	·	·H	·H	·	·	·H	·H	'44
-0.3	·	·	L·	L·	·	·	·	·	·H	·	'40
4.4	·	·	·	L·	·	L·	·	·	·	·H	'48
11.6	·H	L·	·H	·H	·	·H	L·	·	LO·	·H	'64
14.4	·	·H	L·	·H	·	·	L·	LO·	L·	·	'32
17.2	·	·H	LO·	L·	·H	L·	·HI	·	·HI	L·	'36
Counties	G4	H3	G8	E8	G6	E9	H7	G7	I2	E7	

and for those of '28, '40, and '44 to drift upward -- many large residuals in 1936, and almost as many in 1964.

- taking elections in chronological order: nothing exciting (hardly a surprise, since we have just taken out time trends -- a computational slip in how much time trend was taken out of one county showed up in the first draft of this table and has been corrected).

- taking counties from West to East: a weak tendency for '28 to trend up and a strong tendency for '36 to trend down; many large residuals in 1936, almost as many in 1964.

- taking elections from Republican to Democratic: a tendency for G4 (Dawson county) to trend up (to swing wider) and perhaps, for G7 (Hall county) to trend down (to swing less widely).

Of all these appearances, the great activity of 1936 -- combined with a west-east trend in 1936 and substantial activity in 1964 -- seems the most prominent and the most deserving of attention. We will try the next step there.

17D. Looking at the time trends

Before we take another step, however, we owe it to ourselves to look as directly as we can -- not through numbers -- at the fit we have already made. The least that we could do would be to ask: "What do the fitted slopes add to the previous fit?"

Exhibit 18 shows the simplest -- and probably the best -- answer to that specific question. It shows the time trends that were added to the previous fit. (Since we used multiples of a single standard trend in our fitting the additions all vanish at the same date -- half-way from '40 to '44.) It also reminds us that this was a poor question. We ought at least to combine this additional fit with part of the fit we had before.

Exhibit 19 combines the two partial fits that relate to counties:

- the county effect found in exhibit 5 and shown in exhibit 7.
- the time trends found in exhibit 13 and shown in exhibit 18.

We are now able to see quite a lot that we had not noticed before, including:

- the one high -- more Democratic -- county approached the others over these 44 years, but slowly enough to suggest another 100 years before they match.
- a group of counties for which the fit was closely similar in about 1924 tended to spread apart in the ensuing 40 years.

Let us look at the fits for 1924 -- when many lines nearly crossed -- and in 1964 -- as far from 1924 as we can get. Exhibit 20 has the numbers. Since eight of the fitted values are negative, it seems desirable to look at their

exhibit **18** of chapter 17

The Additional Fit -- Time Trends
(expanded vertical scale)

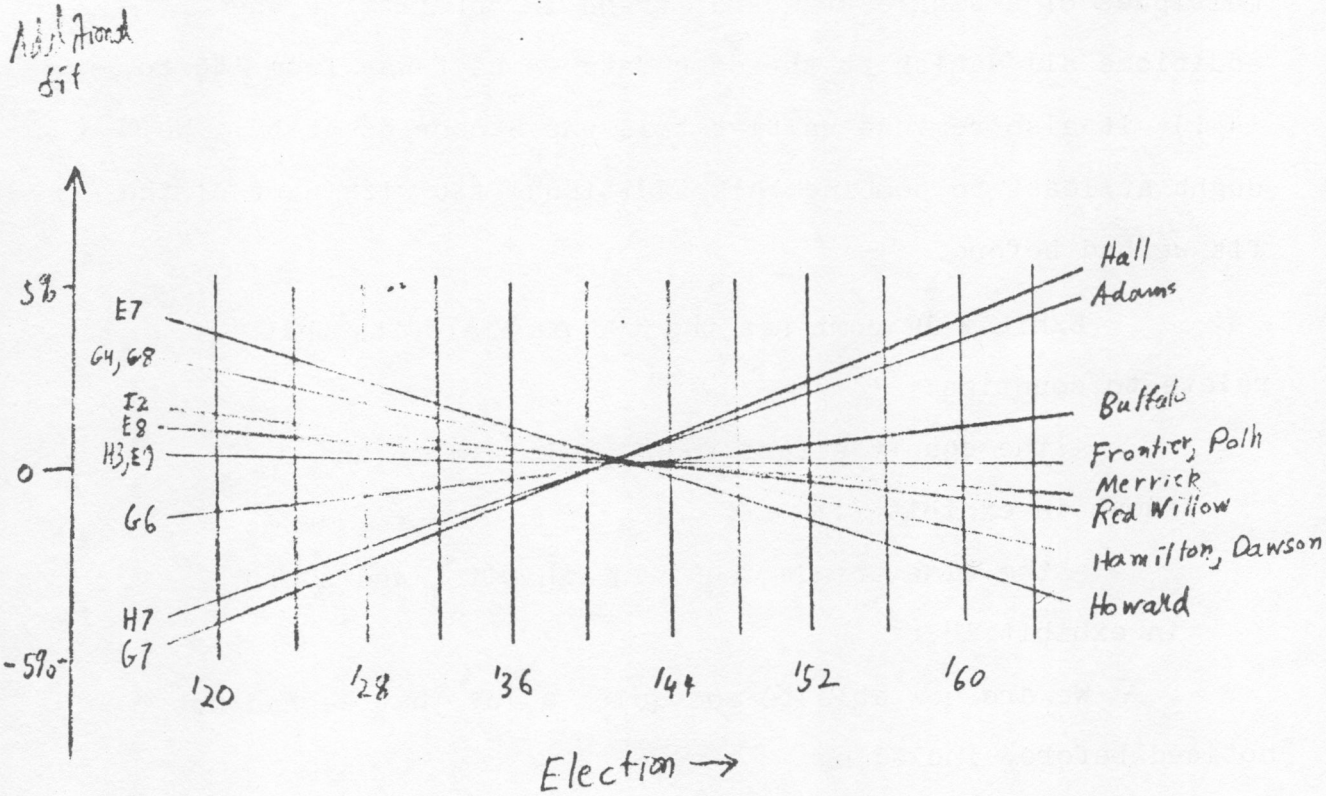


exhibit 19 of chapter 17

Effects and Time-Trends Combined for the 10 Counties

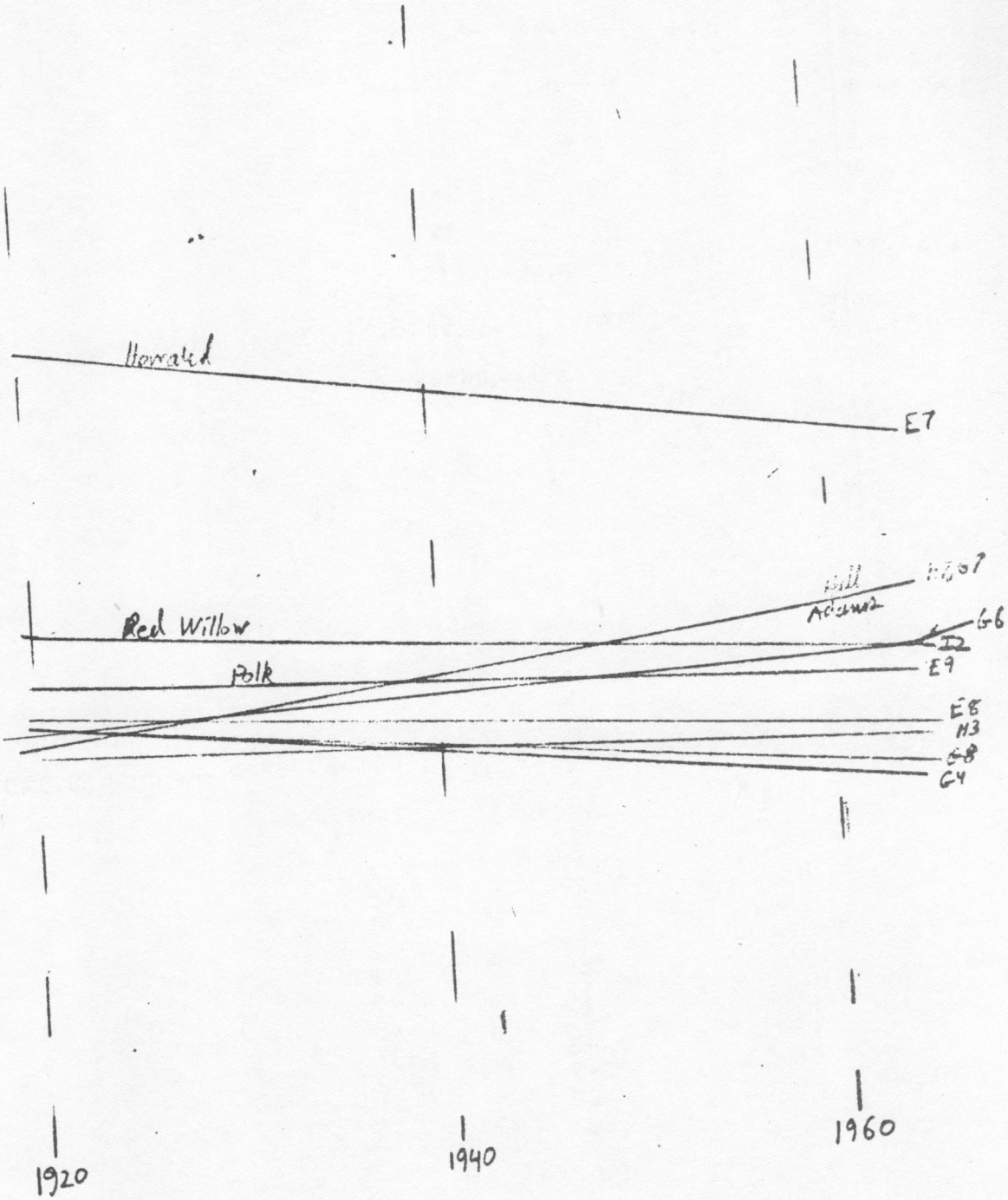


exhibit 20 of chapter 17

County Fits -- Effect and Time-Trend -- in 1924 and 1964

A) 1924 FITS AS A BATCH - unit 0.1%D

```
-4|7
-3|200
-2|246
-1|
-0|5
 0|
 1|
 2|6
 H|H
```

Batch: 10 as 0.1; hinges at 3

Hinges: -3.0 and -0.5

Median: -2.5

B) 1964 FITS AS A BATCH -- unit 0.1%D

```
-7|5
-6|6
-5|
-4|91
-3|
-2|
-1|
-0|27
 0|4
 1|
 2|
 3|
 4|45
 H|H
```

Batch: 10 as 0.1; hinges at 3

Hinges: -49 and +44

Median: -4

C) 1924 FITS RAISED BY 25 and 1964 FITS RAISED BY 4 -- unit 0.1%D

County	1924	1964
I2	53	8
H3	-22	-45
G4	0	-71
G6	1	2
E7	230	141
G7	-12	47
H7	-15	48
E8	-2	-37
G8	-7	-62
E9	20	-3

deviations from a resistant central value, say their median. Panel C of exhibit 20 gives both 1924 and 1964 fitted values referred to their medians. To make sense of these values, we should see them on the map. Exhibit 21 shows the two maps. In 1924, only two counties deviated in fitted values more than ±2.1%D. One was Red Willow (I2) the most southwesterly of our ten, the other of course was Howard (E7), the exceptional case. By 1964, differences in fitted values were more widespread, though there seemed to be a tendency for several pairs of adjacent counties to agree.

* can we see the whole fit? *

Now that we have seen how much more we gained by looking at two parts of the fit together, we feel that we ought to try to look at all three parts at once. This would mean combining:

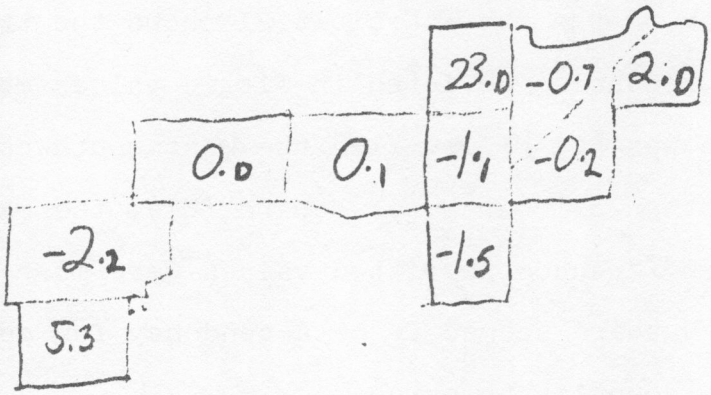
- election effects.
- county effects.
- county time-trends.

(It is a matter of much lower importance whether or not we include the common term, especially since we are likely to be able to add it for no extra cost once we have the others.)

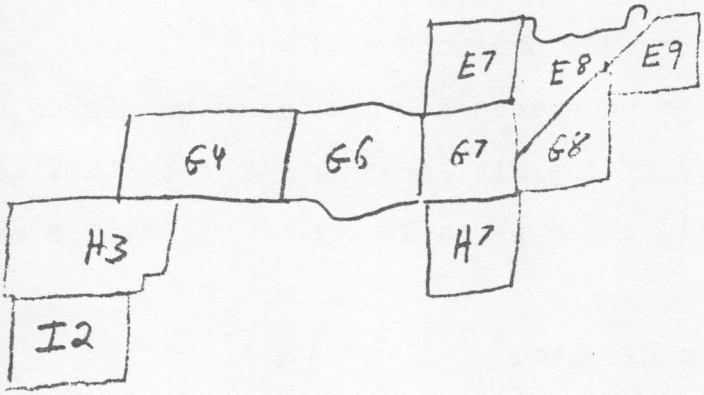
The natural approach to doing this is to start with our straight-line two-way plot -- as in exhibit 7 -- and convert

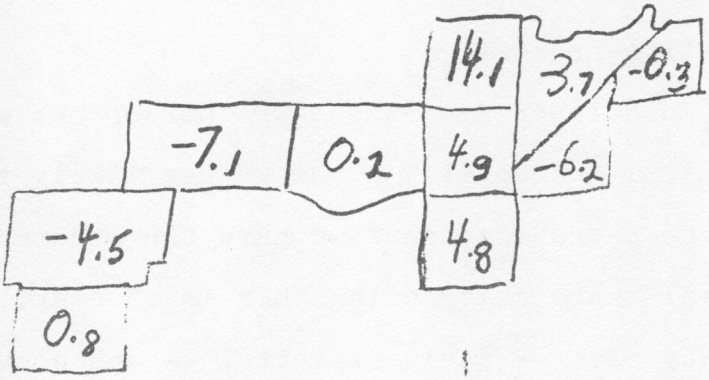
exhibit 21 of chapter 17

Fitted Values in 1924 and 1964 by County
(see exhibit 20)



Fitted values in 1924





Fitted values in 1964

it into a two-way plot showing time trends as well. Two difficulties arise because the trends are linear in time but the two-way plot arranges elections from Republican to Democratic. This means that linear time-trends are very wiggly lines on a two-way plot, as exhibit 22 shows. First:

- a lot of effort must be taken.

There are 9 line segments in the wiggly line, even after we simplify by neglecting two slight changes in slope. In principle this is minor; in practice it could be serious. Second:

- once there are more than a very few counties the plot becomes hard to look at, as exhibit 23 shows.

This is serious. The direct approach cannot be taken effectively.

17E. Bent plots

How are we to avoid this trap? Can we plot all three parts of the fit together effectively at all? When we reach such an impasse by a direct approach, the wise strategy is always to examine our usual assumptions about how things are to be done, seeking for a way to gain by discarding one -- usually a big one.

In a simple two-way plot, as we have used them, there are three sets of lines, actual or implied:

- a set of lines labeled with the versions of one factor, here elections.
- a set of lines labeled with the versions of the other factor, here counties.

exhibit 22 of chapter 17

Linear Time-Trends Plotted Against Election Effects for Two Counties

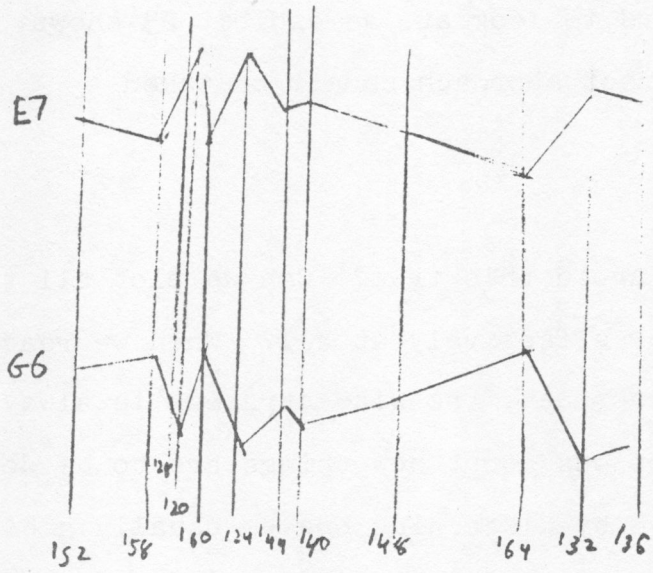
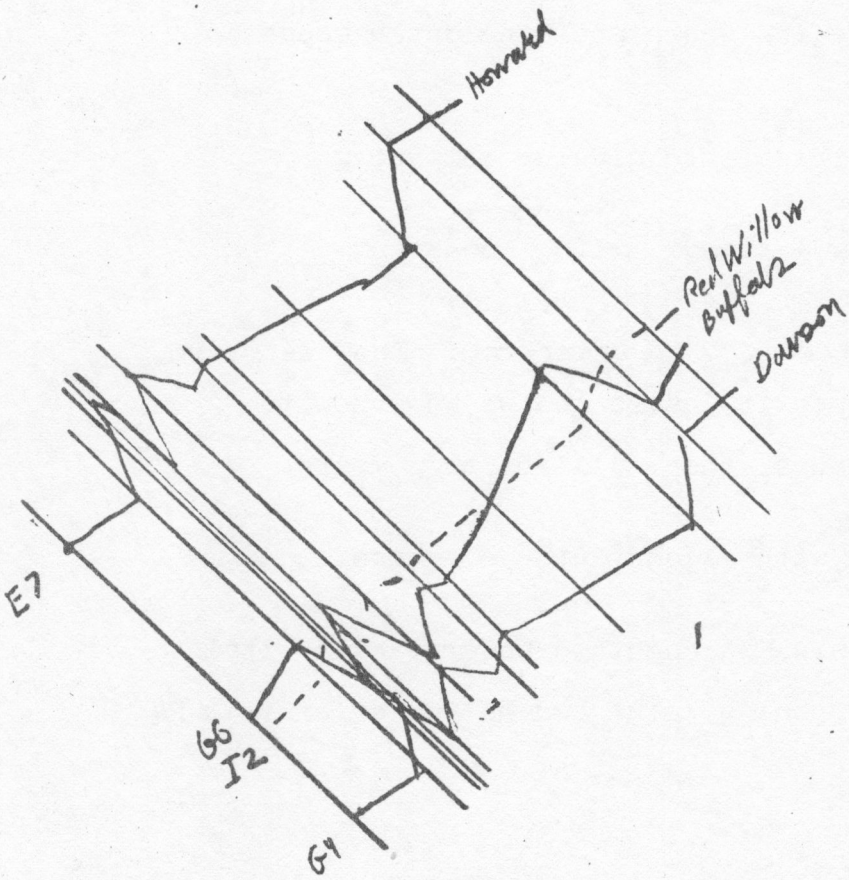


exhibit 23 of chapter 17

The wiggly two-way plot fails, because of confusion

even for only four counties



- a set of horizontal lines, often implicit rather than explicit, labeled (visibly or invisibly) with numerical values of the fit we are describing.

In our example, there must be:

- 12 lines of the first set, one for each election.

- 10 lines of the second set, one for each county.

- enough lines of the third set to guide the eye.

If we are going to deal with all three parts of the fit by making one set of lines wiggly, perhaps it should be the third set, where there need be no more lines than are required to guide the eye and where one line never needs to cross another.

The way to do this is to give up comparing

election + county + trend with constant

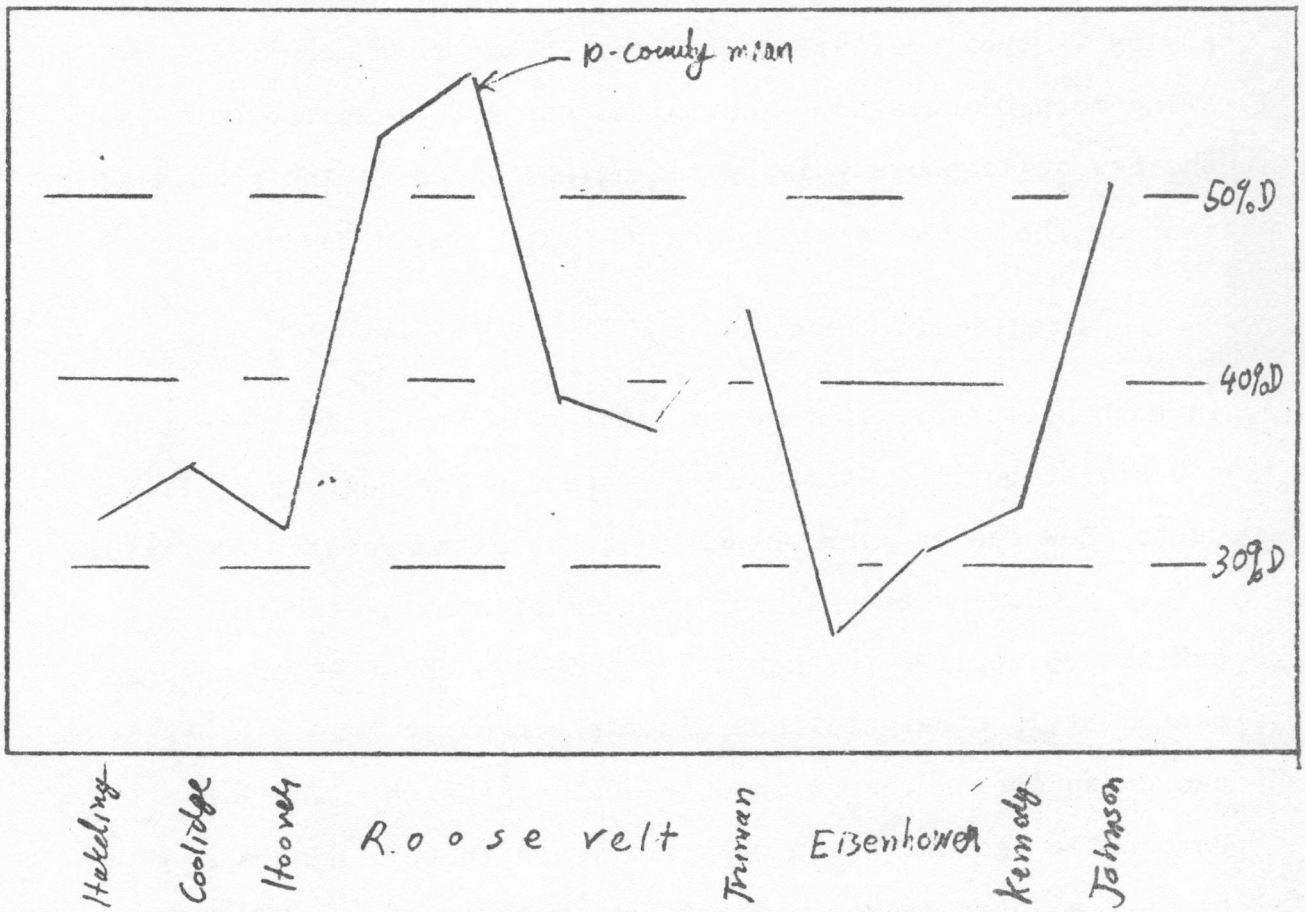
by moving one term from the left to the right. The best term to move is "election" since it leaves two we know how to plot well. Our new patterns, then, will be to compare

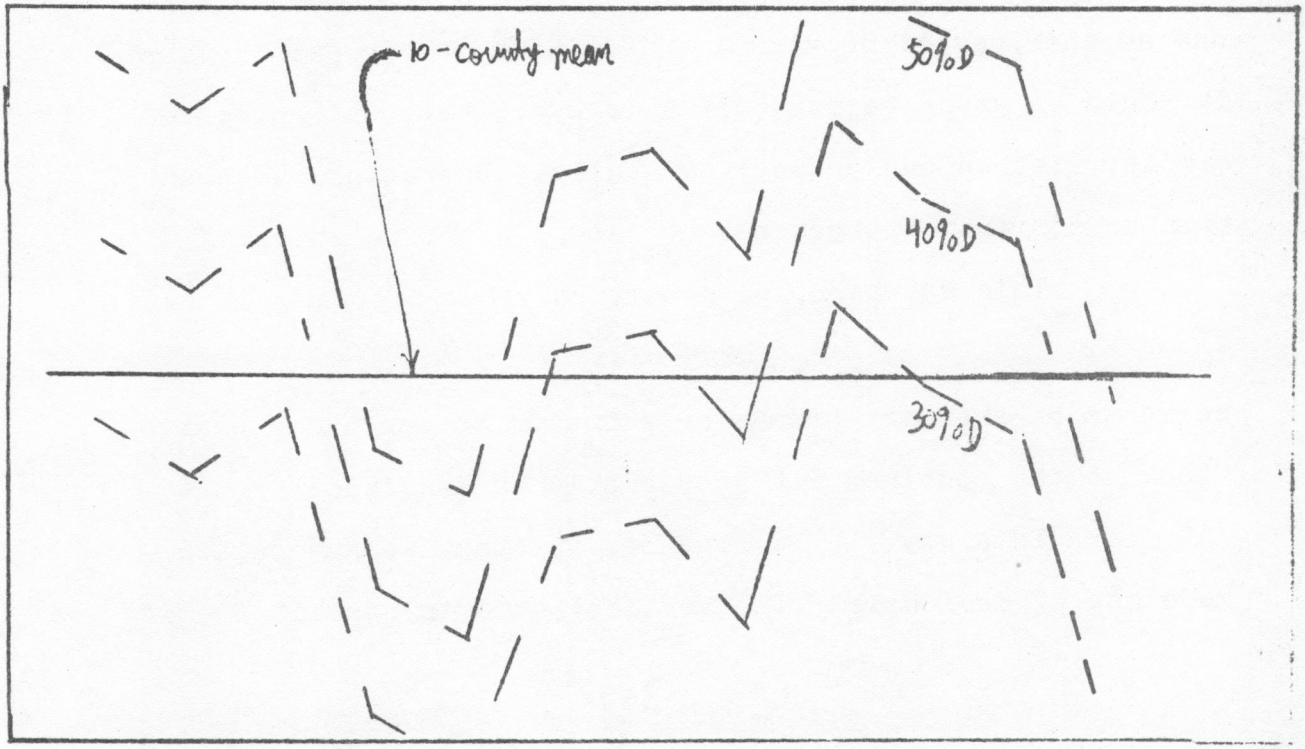
county + trend with constant - election.

Before we do this, we ought to understand a little more visually what is happening. The two halves of exhibit 24

exhibit 24 of chapter 17

Two Ways of Charting the Ebb and Flow of Politics





show the election means for the 10 counties in the two styles -- the familiar and the new. When these counties swing more Democratic -- as in the early Roosevelt elections, the new style shows this by the dashed lines moving down, rather than by the solid lines moving up. The result is the same

ten-county mean above 50%D in 1932 and 1936 only

in both pictures. It takes a little time to get used to the new style, but this time is well-spent. The new style allows us to show easily many things that would otherwise be difficult.

Coming back to our problem of showing the full fit, exhibit 25 applies the new style. The 10 counties are represented by 10 straight lines, just as in exhibit 19. Lines can cross one another without visual confusion. The % Democratic against which we are to read these lines is shown by half a dozen broken wiggly lines which move parallel to one another, avoiding visual confusion by their parallelism. It would be wrong to say that this exhibit, as a whole, looks simple. Yet we can grasp it visually, separating the county time trends from the general ebb and flow of politics.

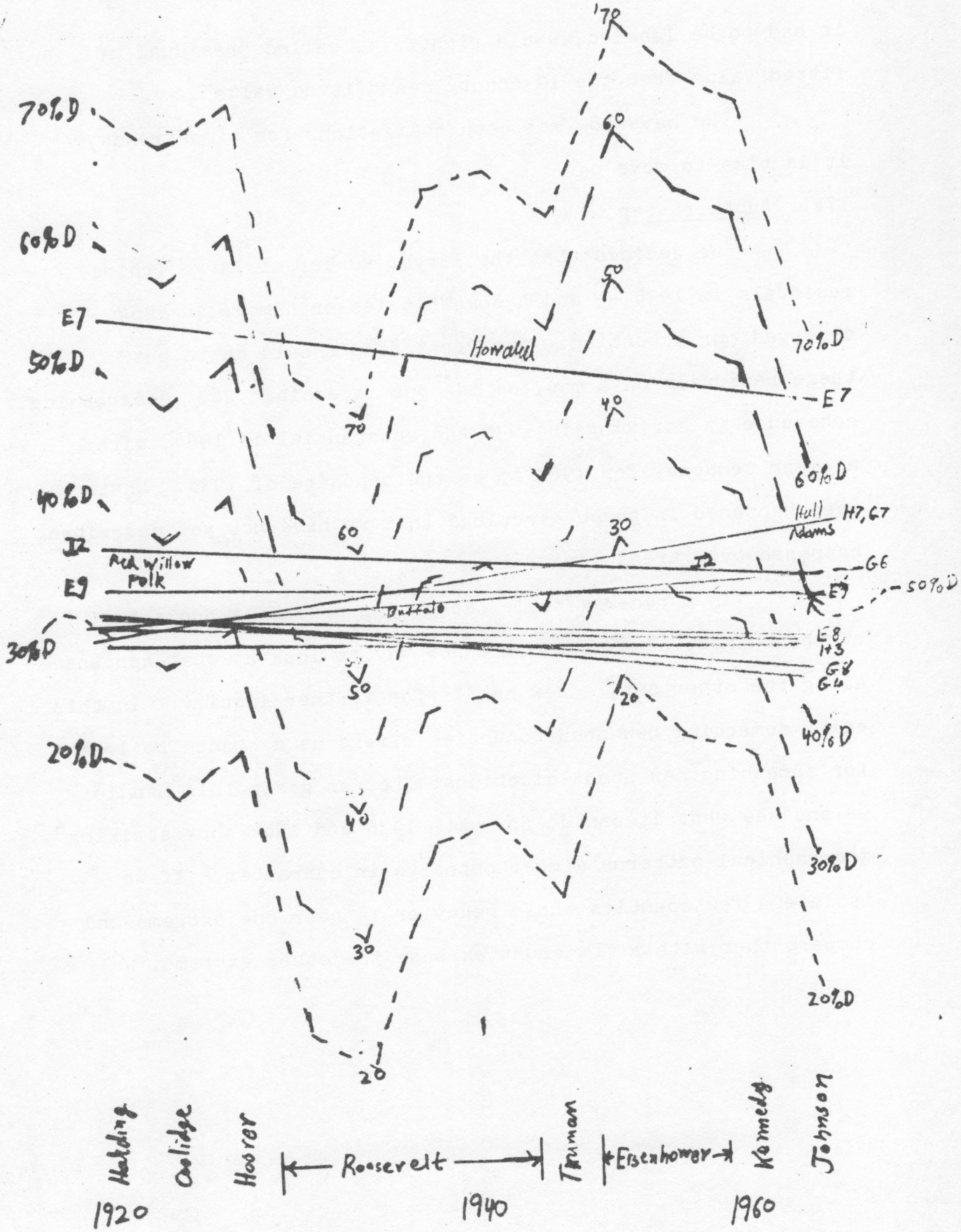
This approach, naturally called a

bent plot

seems to be the most promising solution to our basic charge: "Look at the combined fit as directly as possible!"

In a way, it is right to say that we have again made use of residuals. The vertical axis in exhibit 25, if

10 Nebraska Counties against the political ebb and flow
(fitted values, see exhibits 5 and 19)



it had to be labeled, would rightly be called "residual of fitted value from the 10-county mean fitted value".

We have now met both obligations for time-trends, it is time to move on.

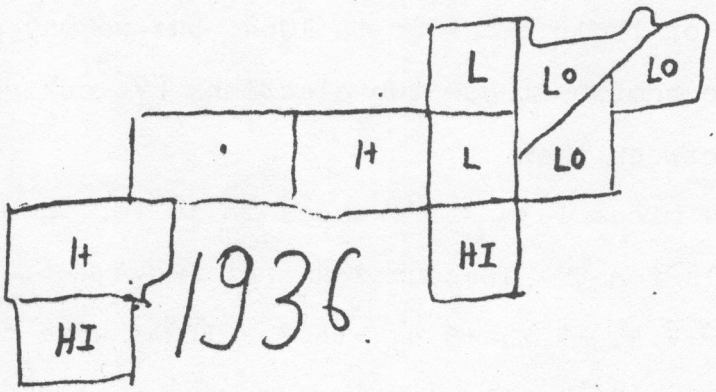
17F. Another step

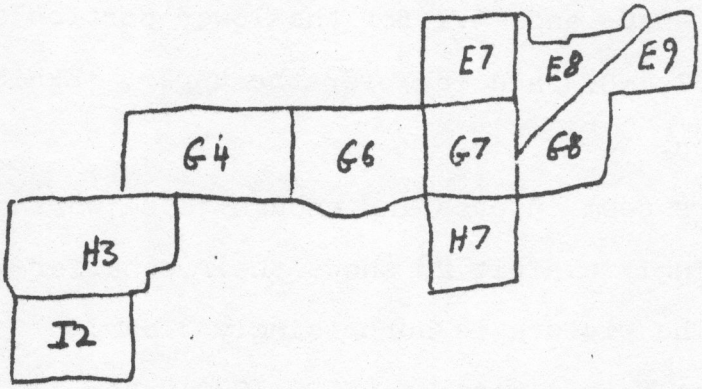
We decided that the larger number of large (third) residuals in 1936 -- and a somewhat lesser number in 1964 -- deserved our attention. Our first step should be to put these residuals on a map, as is done in exhibit 26. Geographical coherence is striking in 1936 and substantial in 1964, with a clear tendency for 1964 to be the opposite of 1936. Whatever happened in these elections that we have not yet described, happened over areas larger than counties.

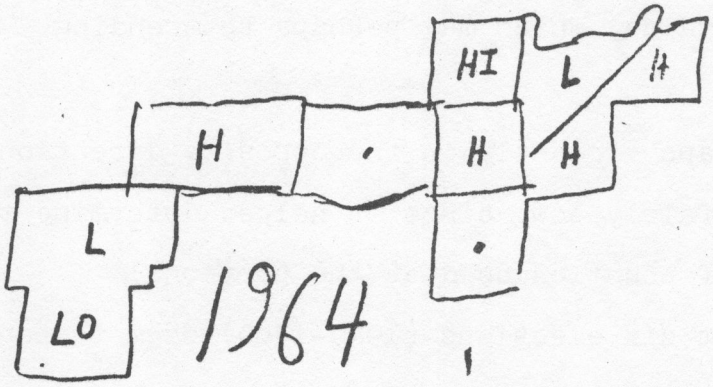
In a two-way table, something new about one set of stripes almost always offers a new way to look at what happens along the other set, a new handle for further inquiry. In this case, something new about counties offers us a chance to look for something new about elections. Let us grasp this handle -- and see what it can do for us. 1936 and 1964 show striking geographical patterns almost opposite in cahracter. If we isolate a few counties whose behavior is near one extreme and compare them with a few counties near the other extreme, we

exhibit 26 of chapter 17

Third Residuals in 1936 and 1964
(see exhibit 17)







can have a new measure of elections. (One that is sure to have 1936 and 1964 close to opposite extremes.)

We could select these counties in terms of their behavior in 1936, or their behavior in 1964, but we may do still better if we combine these two elections by looking at the differences between them.

The 1964 minus 1936 differences in %D for the detrended (= third) residuals for the 10 counties are: -8.4, -6.0, -5.1, -2.4, -2.3, +2.2, +4.4, +4.7, +5.4, +6.7. If we were to seek out a tilt for each individual election, we ought to take the counties with -8.4, -6.0, and -5.1 for the lower portion and those with +4.4, +4.7, +5.4 and +6.7 for the upper. Exhibit 27 shows the calculation.

The numbers seem interesting enough to be worth looking at. Accordingly exhibit 28 shows their behavior as a function of time. The picture is surprisingly clear:

- the first four elections (1920-1932) show a rather stable value, which may perhaps be trending upward.
- 1936 snaps violently in the opposite direction (it will be unfairly low, since it helped determine the extreme sets of counties used as the portions).
- the next six elections (1940-1960) seem to be relatively consistent, though there seems to be an upward trend. Their level is appreciably lower than that of the first four elections.

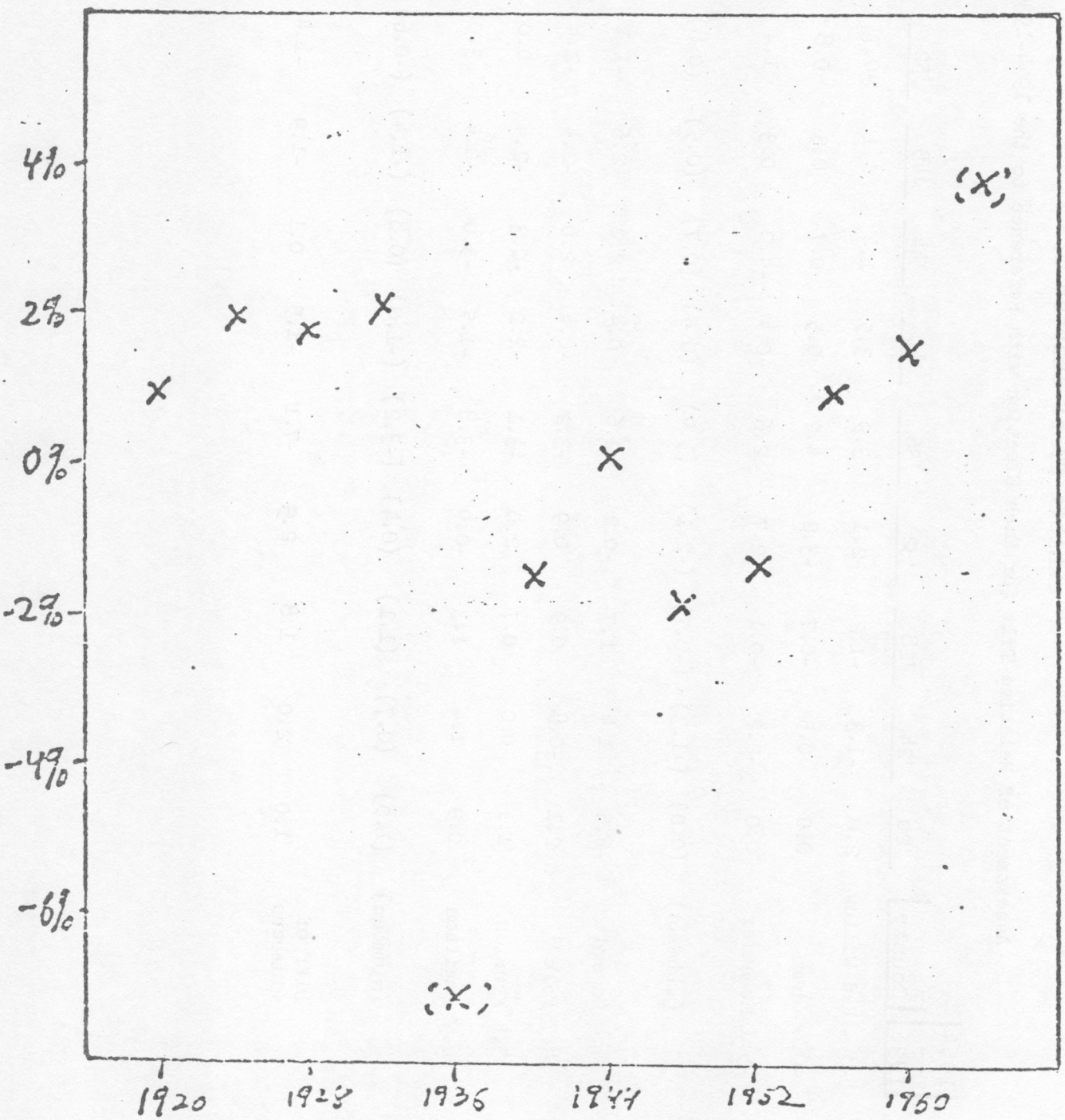
exhibit 27 of chapter 17

Assessment of Relative Tilt for Each Election with Reference to the 1964-1936 Difference

1964 -- 1936	County	'20	'24	'28	'32	'36	'40	'44	'48	'52	'56	'60	'64
-8.4	Red Willow	2.0	-1.3	-1.2	-2.1	3.9	3.5	1.4	1.1	-0.7	-1.9	-0.6	-4.5
-6.0	Adams	0.0	0.5	-0.7	-3.2	4.2	0.5	0.7	0.5	0.8	-0.7	-1.2	-1.8
-5.1	Frontier	0.0	-2.2	-0.1	2.7	2.6	-0.4	-4.0	0.3	1.4	2.4	-1.7	-2.5
	(midmean)	(0.0)	(-1.3)	(-0.7)	(-2.1)	(3.9)	(0.5)	(0.7)	(0.5)	(0.8)	(-0.7)	(-1.2)	(-2.5)
+4.4	Howard	-6.5	1.4	1.7	-0.2	-2.6	0.4	2.2	2.6	-1.1	+0.2	+0.1	+1.8
+4.7	Polk	2.1	-0.8	0.9	0.9	-3.2	-0.4	2.0	-2.4	-2.4	1.2	0.7	1.5
+5.4	Merrick	1.1	0.0	0.7	2.0	-3.1	-2.2	-0.3	-2.4	0.0	0.2	1.9	2.3
+6.7	Hamilton	0.9	1.4	1.3	-0.9	-3.8	-1.5	-3.0	-0.4	3.0	-0.1	+0.4	2.9
	(midmean)	(1.0)	(0.7)	(1.1)	(0.4)	(-3.2)	(-1.0)	(0.8)	(-1.4)	(-0.6)	(0.2)	(0.6)	(2.0)
	Diff of midmeans	1.0	2.0	1.8	2.5	-7.1	-1.5	0.1	-1.9	-1.4	0.9	1.8	4.5

exhibit 28 of chapter 17

Difference of midmeans of detrended residuals
4 Northerly counties (1964-1936 >> 0)
referred to
3 Southerly counties (1964-1936 << 0)



- 1964 definitely either jumps up or continues an upward trend (it, too, will be unfairly extreme because of its use in choosing the portions).

We have found another clue. It seems likely we are at the bottom of the bag. But we may as well take another look. We have just identified two sets of relatively peaceful elections -- 1920-1932 and, say 1940-1952 -- as differing in behavior. How does this appear in terms of individual counties? Exhibit 29 shows the calculation and exhibit 30 the geographical pattern. So far as these elections are concerned, there seems to be a dividing line between the three Eastern counties and the other seven.

We need to stop and give ourselves a word of caution now. When we fitted time trends to our (second) residuals (in exhibit 13), we compared 1920-1932 with 1952-1964, choosing these as early and late. After making that adjustment, we have compared 1920-1932 with 1940-1952 because these seemed to be quiet times. Surely there must be some overlap, some tendency for the first fitting to try to make part of the adjustment that could have been made by the second one. This is a reasonable warning -- we all ought to heed it by caution in making interpretations, but we shall do nothing more explicit about it here.

There is another step we could take. This is to clear our trail in another way by eliminating the two most

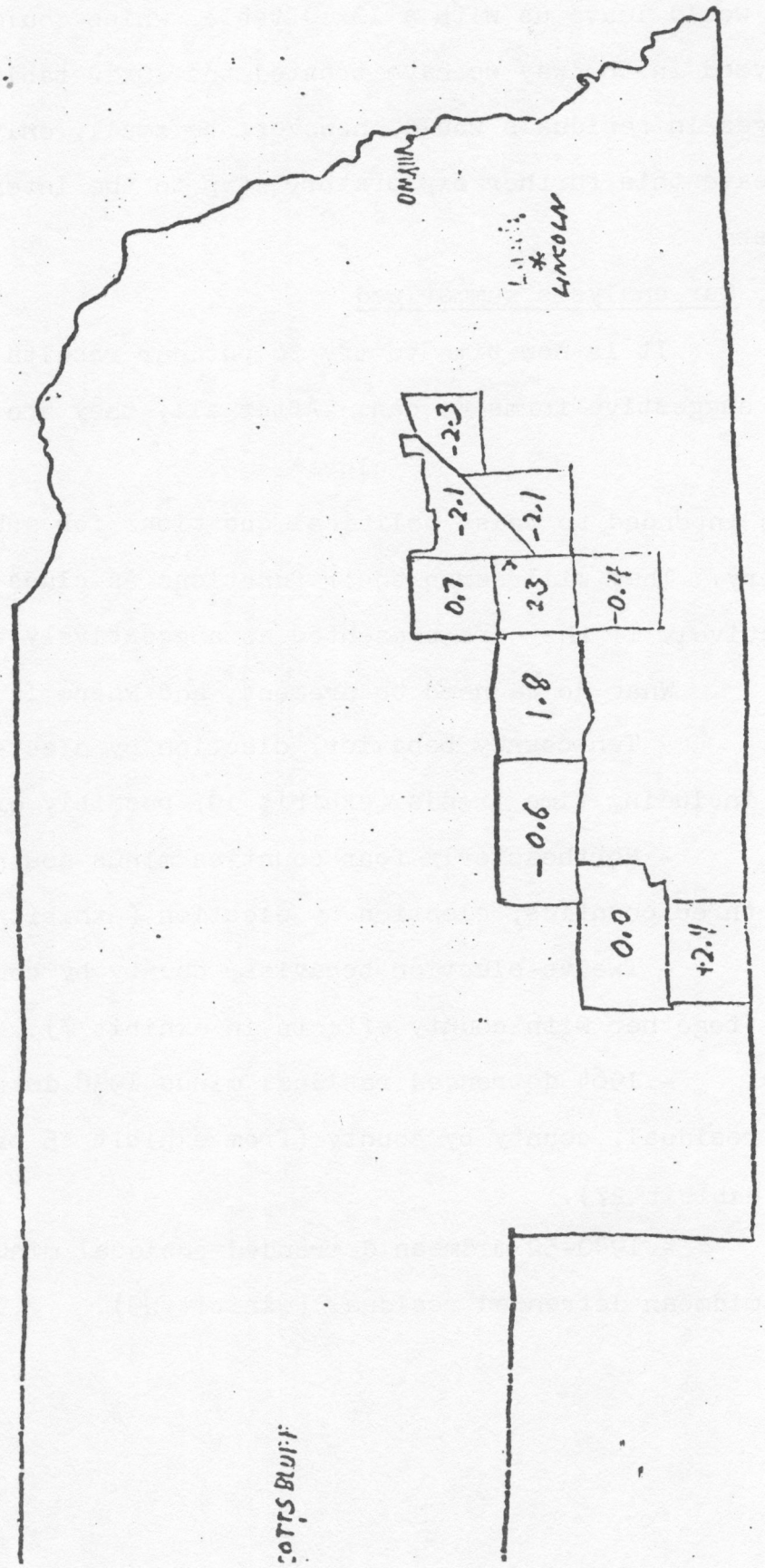
exhibit 29 of chapter 17

Comparison of Early Quiet and Late Quiet Elections

County	Midmean 1920-1932	Midmean 1940-1952	Difference
Adams	-0.4	0.6	1.0
Buffalo	-0.7	1.1	1.8
Dawson	0.5	-0.1	-0.6
Frontier	0.0	0.0	0.0
Hall	-2.6	-0.3	2.3
Hamilton	1.1	-1.0	-2.1
Howard	0.6	1.3	0.7
Merrick	0.9	-1.2	-2.1
Polk	0.9	-1.4	-2.3
Red Willow	-1.2	1.2	2.4

exhibit 30 of chapter 17

Location of counties in Nebraska and results of comparing later quiet elections (1940-1952) with earlier ones (1920-1932) as in Table 19



unusual elections, 1964 and 1936, from all consideration. This would leave us with a 10x10 table, which could then be analyzed in the way we have treated the 10x12 table. The changes in residuals would, however, be small, only up to ±0.2. We leave this further exploratory step to the interested reader.

17G. Our analysis summarized

It is now time to try to put our results into the most suggestive forms we can. After all, they are only clues intended to raise political questions for subject-matter inquiry. They will serve their functions as clues most effectively if they are presented as suggestively as possible.

What do we need to present, and where is the raw material?

- Ten-county behavior, election by election including time trends (exhibit 19, possibly exhibit 25).
- Northeasterly four counties minus southwesterly three counties, election by election (exhibit 28).
- Twelve-election behavior, county by county (together with county effects in exhibit 7).
- 1964 detrended residual minus 1936 detrended residual, county by county (from exhibit 15 or exhibit 27).
- 1940-52 midmean detrended residual minus 1920-32 midmean detrended residual (exhibit 29).

- Relation of time trends to vote growth (exhibit 14).
- The unusual behavior of Howard county in 1920.

Where one can, the way to be suggestive is to be graphical. So let us try to summarize all these items graphically. Exhibit 31 shows the overall behavior for the ten counties, both alone and in contrast with all of Nebraska and all the U.S. Making this graph do triple duty has reduced its impact, but we do have our attention called to such facts as:

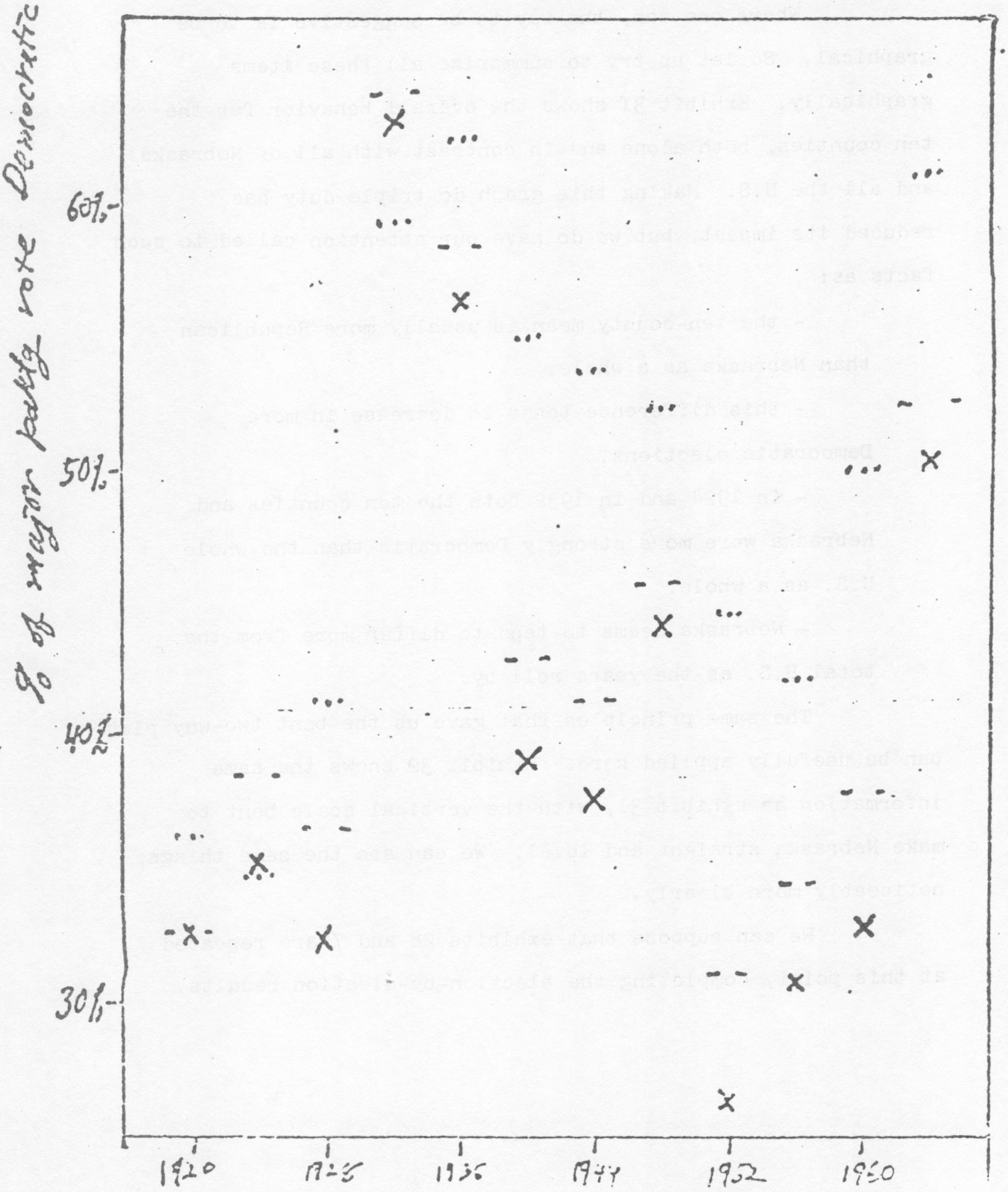
- the ten-county mean is usually more Republican than Nebraska as a whole.
- this difference tends to decrease in more Democratic elections.
- in 1924 and in 1932 both the ten counties and Nebraska were more strongly Democratic than the whole U.S. as a whole.
- Nebraska seems to tend to differ more from the total U.S. as the years roll by.

The same principles that gave us the bent two-way plot can be usefully applied here. Exhibit 32 shows the same information as exhibit 31, with the vertical scale bent to make Nebraska straight and level. We can see the same things, noticeably more clearly.

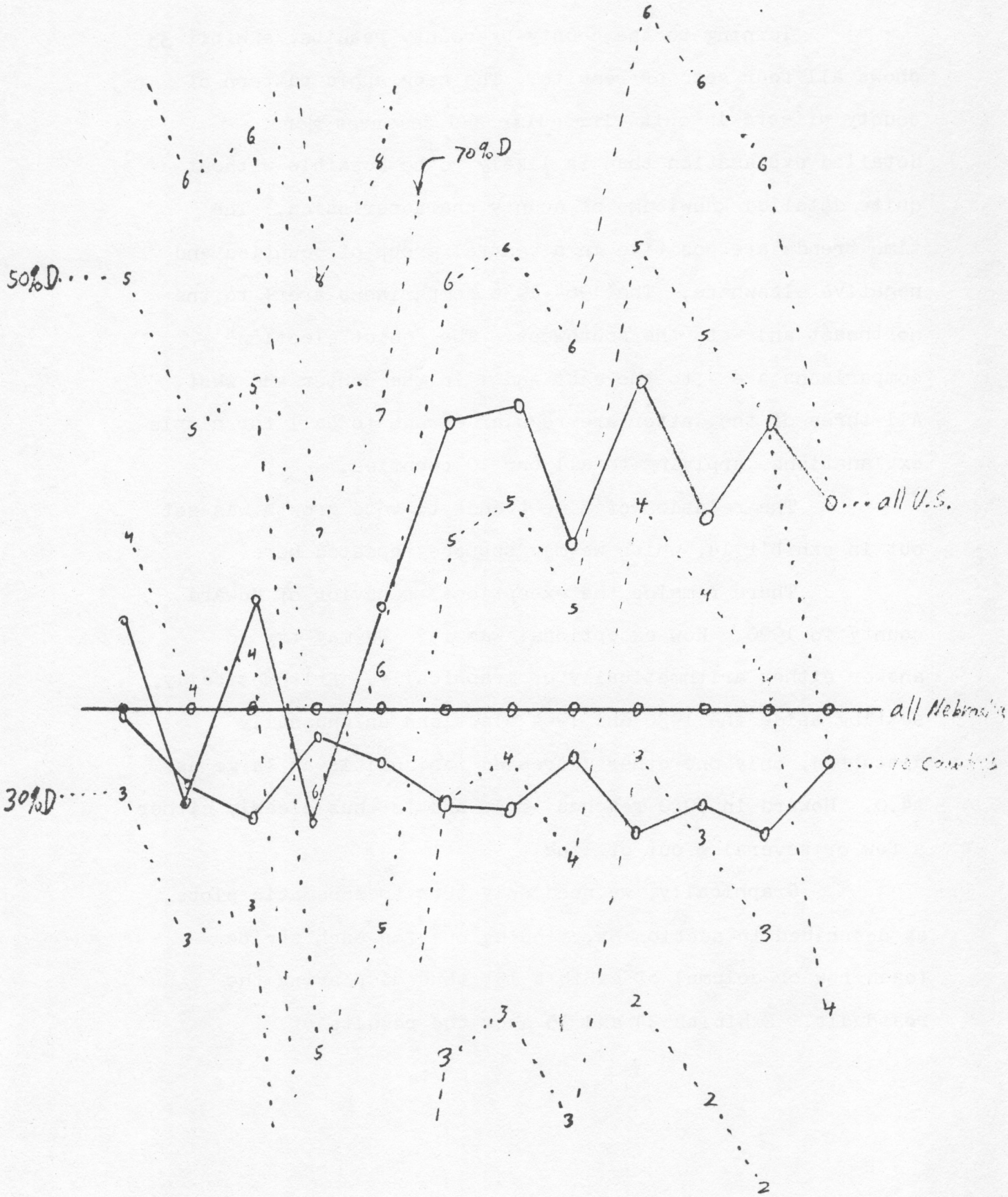
We can suppose that exhibits 28 and 7 are repeated at this point, completing the election-by-election results.

exhibit 31 of chapter 17

Behavior of the ten counties through time
x = ten counties
- - = all Nebraska
... = all U.S.



The same information referred to Nebraska



Turning to the county-by-county results, exhibit 33 shows all four sets of results. The geographic pattern of county effects is quite irregular and deserves more detailed explanation than is likely to be possible without quite detailed knowledge of county characteristics. The time trends are positive in a central group of counties and negative elsewhere. The 1964-1936 comparisons are + to the northeast and - to the southwest. The "quiet election" comparisons are -, to the east and + in the center and west. All three of the latter are regular enough to call for simple explanations, applying to all our 10 counties.

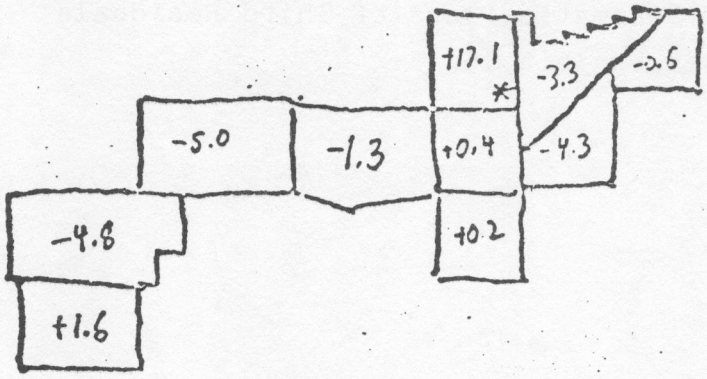
The relation of time trends to vote growth was set out in exhibit 14, which we may suppose repeated here.

There remains the exceptional behavior of Howard county in 1920. How exceptional was it? We may try to answer either arithmetically or graphically. Arithmetically, setting aside the 1936 and 1964 elections as unusually disturbed, only one other detrended residual is as large as ±4.0. Howard in 1920 reaches -6.5, and is thus clearly either a few or several % out of line.

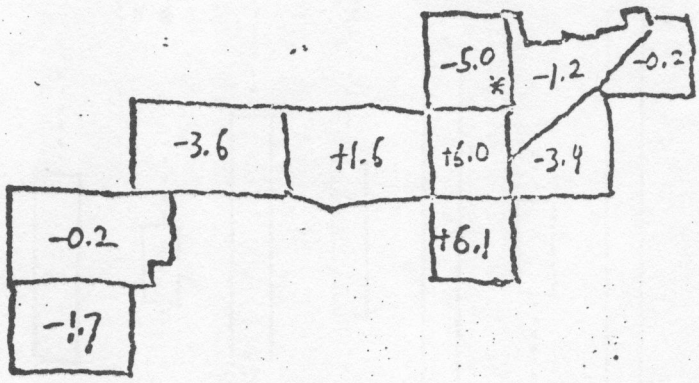
Graphically, we need only turn to schematic plots, as described in section 5D, finding one for each stripe (each row or column) of exhibit 15, thus displaying the residuals. Exhibits 34 and 35 show the results of

exhibit 33 of chapter 17

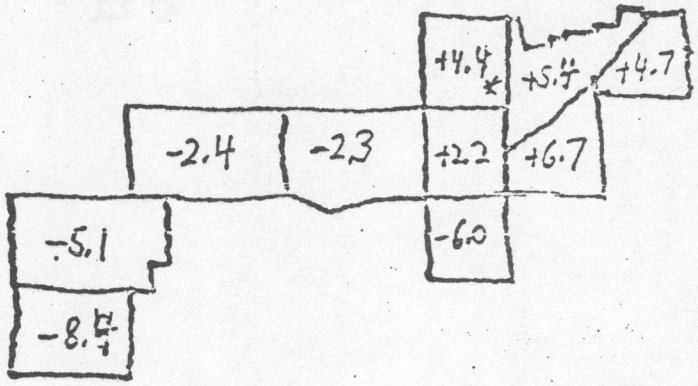
County to county comparisons



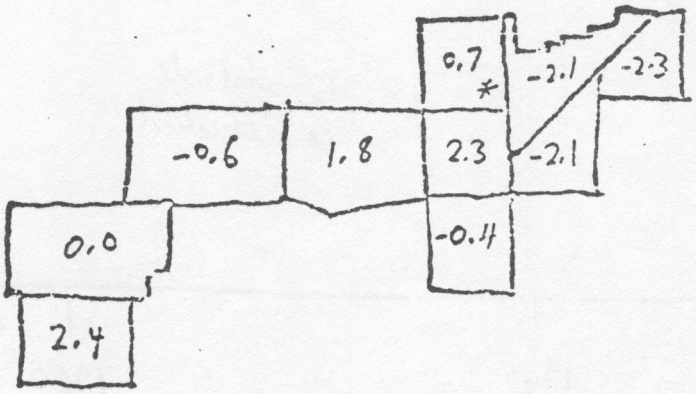
County effects
(+ = more Dem)



Time Trend in
Second residuals
(+ = toward Dem)



1964 compared with
1936 in terms of
detrended residuals
(+=later more Dem)



Quiet elections (1945-1955)
compared to
quiet elections (1920-1932)
in terms of detrended
residuals.
(+ = later more Dem)

* Indicates Humphrey county

exhibit 34 of chapter 17

Election-by-Election Schematic Plots of Third Residuals

(see exhibit 15)

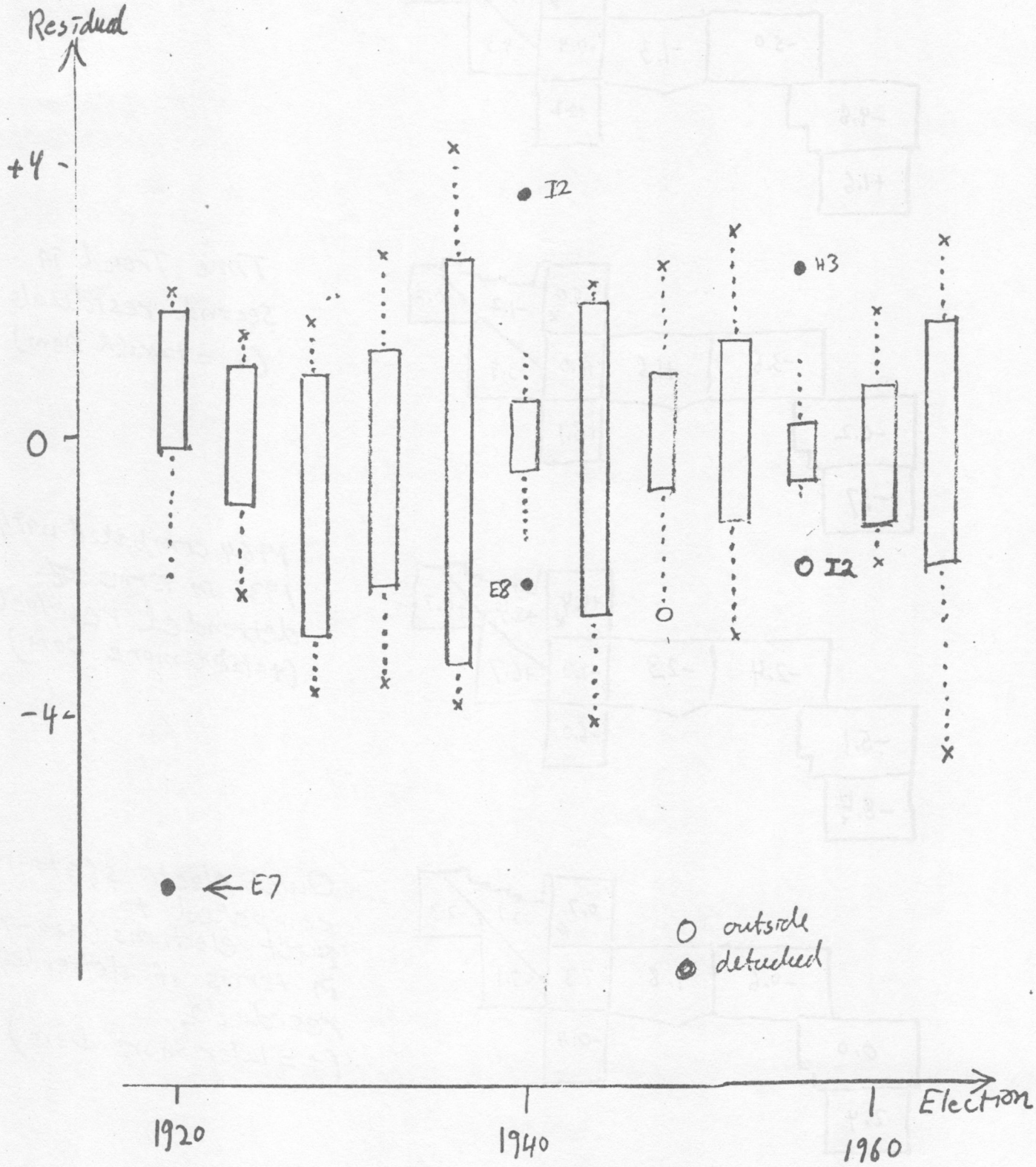
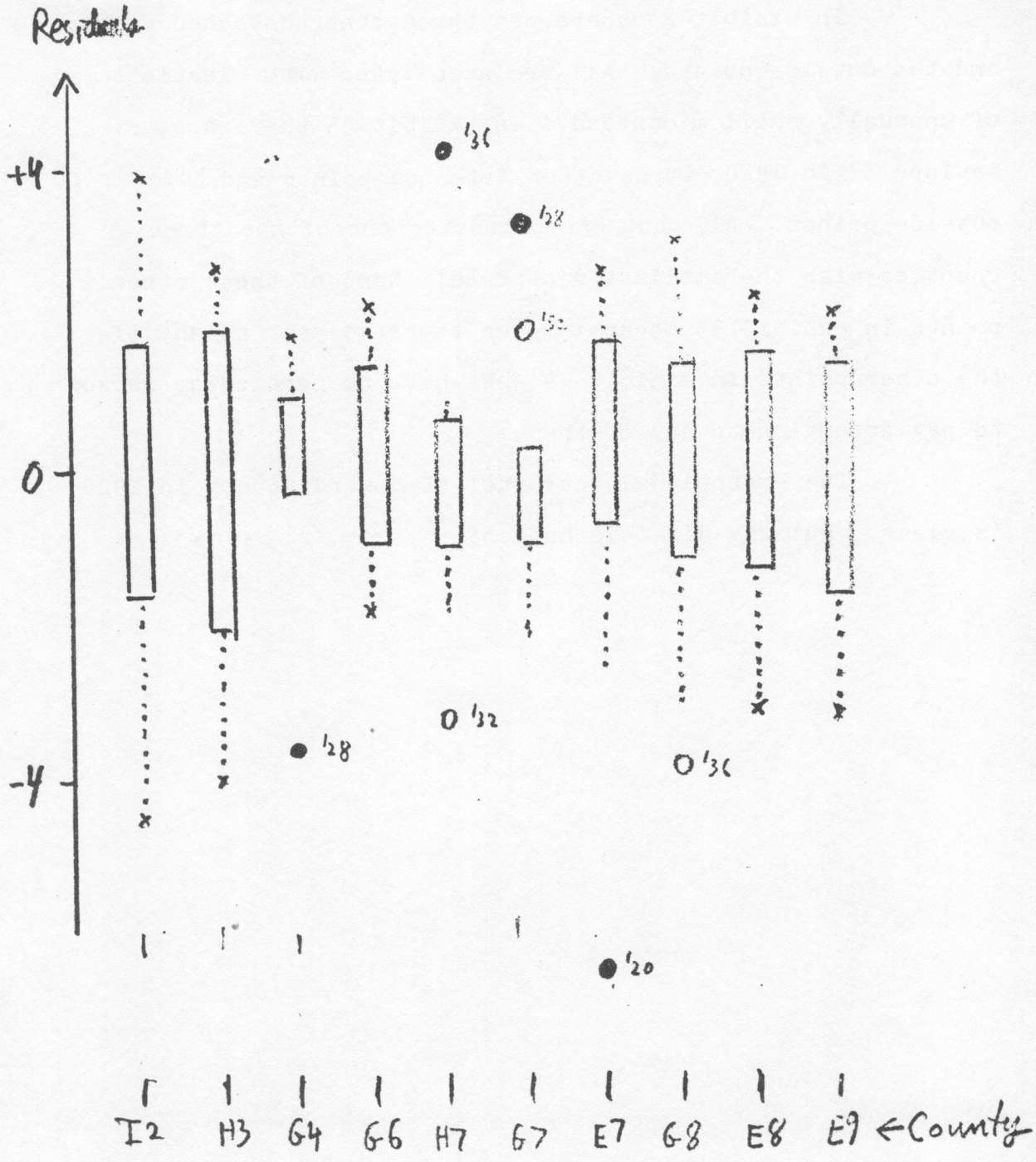


exhibit 35 of chapter 17

County-by-County Schematic Plots of Third Residuals
(see exhibit 15)



such an approach applied first to the 12 elections and then to the 10 counties. The clearly exceptional character of E7 (Howard County) in 1920 is shown in both plots. It is detached both from the other counties in 1920 and from the other elections in Howard County.

In exhibit 34 there are three other detached points and two outside points. All are associated with elections of unusually small midspread. In exhibit 35 there are, besides E7 in 1920, three other detached points and three outside points. All but one belong to one of the three counties with the smallest midspread. None of these other points in exhibit 35 occur in even the same year as any of the other points in exhibit 34. We have no particular reason to pay attention to any of them.

The exceptional character of Howard County in 1920 is clear. But why did this happen?

*****

17H. How far have we come?

Seven sections of this chapter, and 35 plots and tables have been devoted to one example of a two-way table of responses. That there might be this much in one 10x12 table may have surprised you. However, this sort of situation, in which layer after layer can be stripped off, is more nearly the rule than the exception. Once we acquire the structure of a two-way table, there are many things we can do -- and it is often important to do several of them in a particular instance.

We have seen examples of a number of specific techniques:

- of the peeling away of row and column means or mean residuals to reach one form of elementary analysis.
- of the use of coded tables of residuals to show to eye and brain what tables of numbers can tend to conceal.
- of the recognition and removal of trends as well as means.
- of how something recognized along one way (in response to one factor) offers us a way to ask more sharply about the apparent effect along the other way (in response to the other factor).
- of how outside information, such as the geographic arrangement of counties, can contribute to how believable (how clearly not accidental) our results are to be.

The general message is clear: Once the data has more than the simplest structure, there is much to be gained from successive layers of "clearing the trail" by peeling off one piece of partial description after another. Two-way data is much more challenging than one-way (or no-way) data and is likely to be correspondingly more rewarding.

To keep our analyses as clear as we could for those not yet at home with two-way analyses, we have used means in our analyses and avoided median polish. This means that we have made the arithmetic harder than necessary and have accepted the risk that an analysis by means may not be good enough (because a resistant analysis is needed. The next chapter will fill in these gaps.

****

17P. Problems for chapter 17

1. Panel A of exhibit 36 gives presidential voting % for the 10 western counties of Nebraska. Carry out a thorough analysis.

2. Do the same for the 5 counties of Rhode Island in panel B.

3. Do the same for the 8 counties of Connecticut in panel C.

4. Do as much the same as you can for the 3 counties of Delaware in panel D.

5. Exhibit 37 gives the local times of moonrise for the first 14 days of July 1971 at various latitudes. Make a two-way analysis, including plots. Continue as far as you can.

6. H. A. Schuck (1949. Relationship of catch to change in population size of New England haddock. Biometrics 5: 213-231. His table 2 on page 229) gives estimates of the relative population of each age of George's Bank haddock in terms of numbers caught per day for ages from 1 to 8 years and years from 1931 to 1947. If we wanted to make a 2-way analysis of such data, what would it seem sensible to use as factors? Why? How would you expect to express the response? Once you have clearly and firmly answered these questions, select from the following data a 6-by-6 table and analyze it. What do you see in the residuals? See exhibit 38 for data.

(The next 5 problems can be shared among a group of friends.)

7. Kaye and Otis (1948. Isomeric xylene and ethylbenzene mixtures. Analytical Chemistry 20: 1006-1008. Their table I on page 1007) give absorption coefficients (0.1-millimeter cell) for concentrations of (0.0476, 0.0238, 0.0119, 0.0060, 0.0030, 0.0015) milliliters of para-xylene per milliliter of carbon disulfide and the following wave lengths: At 12.54 microns (19.51, 21.12, 22.10, 22.50, 22.60, 24.90). At 12.98 microns (0.29, 0.42, 0.67, 1.00, 1.33, 2.00). At 13.44 microns (0.15, 0.17, 0.25, 0.50, 1.00, 2.00). At 14.30 microns (0.11, 0.13, 0.24, 0.50, 1.00, 2.00). Make a reasonable two-way analysis, including plots.

8. Kaye and Otis (see above) gave data for meta-xylene at the same concentrations, namely: At 12.54 microns (0.23, 0.29, 0.42, 0.66, 1.00, 1.33). At 12.98 microns (15.97, 19.83, 20.92, 21.66, 22.33, 24.66). At 13.44 microns (0.23, 0.29, 0.34, 0.50, 0.66, 1.33). At 14.30 microns (1.11, 1.21, 1.18, 1.50, 1.66, 2.00). Do the same.

9. Kaye and Otis (see above) also gave data for ortho-xylene at the same concentrations, namely: At 12.54 microns (0.15, 0.17, 0.17, 0.33, 0.66, 1.33). At 12.98 microns (0.42, 0.50, 0.68, 1.00, 1.66, 2.66). At 13.44 microns (25.65, 31.09, 33.36, 35.50, 36.33, 38.00). At 14.80 microns (0.15, 0.17, 0.17, 0.33, 0.66, 1.33). Do the same.

10. Kaye and Otis (see above) also gave data for ethylbenzene at the same concentrations, namely: At 12.54 microns (0.90, 0.97, 1.09, 1.16, 1.33, 2.00). At 12.98 microns (1.76, 1.89, 2.01, 2.33, 3.00, 4.00). At 13.44 microns (4.62, 4.78, 4.96, 5.50, 5.66, 6.00). At 14.30 microns (11.34, 11.84, 12.10, 12.66, 13.00, 13.30). Do the same.

11. Combine and compare the analyses of the last four problems. What further calculations or plots are needed? Make them.

exhibit 36 of chapter 17 -- more voting

%D of major party Presidential vote in 10 western Nebraska counties, and in Rhode Island, Connecticut and Delaware

A) 10 WESTERN NEBRASKA counties

	County	Election											
		'20		'28		'36		'44		'52		'60	
B1	Sioux	28.7	23.7	27.0	60.1	58.7	36.6	31.8	44.2	21.5	34.3	35.8	35.2
D1	Scott's Bluff	23.3	24.9	17.4	53.8	58.7	35.8	35.0	44.8	27.1	40.1	39.3	47.8
F1	Bameer	21.1	26.4	12.9	55.6	57.0	33.4	28.9	42.6	22.3	36.0	33.2	35.4
G1	Kimball	27.1	25.2	25.3	61.5	57.5	38.2	33.0	39.6	22.4	29.2	32.1	44.1
A2	Danes	33.3	27.4	26.4	54.0	57.2	35.7	34.5	38.5	24.4	26.1	29.2	38.4
C2	Box Butte	31.7	35.1	29.0	60.3	62.9	42.7	36.7	46.3	21.9	31.3	37.4	41.9
E2	Morrill	32.8	38.9	24.6	58.8	59.6	42.2	35.7	47.2	26.5	36.2	37.2	42.7
G2	Cheyenne	24.5	24.4	37.4	70.5	68.2	44.8	39.6	49.7	34.5	36.4	40.4	46.2
B3	Sheridan	31.1	30.5	28.8	61.8	56.6	33.1	28.6	35.1	19.3	25.4	28.5	32.1
E3	Gorden	31.3	38.8	21.6	62.0	49.7	38.5	36.6	36.0	19.0	32.4	23.6	33.6
G3	Denel	31.9	29.0	25.2	63.4	57.7	32.8	26.5	28.3	14.9	20.2	23.8	35.4
B)	The 5 counties of RHODE ISLAND												
	Bristol	29.8	38.0	51.9	55.5	52.3	52.9	56.1	58.6	49.5	40.1	59.6	76.2
	Kent	28.6	32.8	39.4	48.4	49.4	49.2	50.6	48.4	32.1	34.1	55.7	78.3
	Newport	19.3	29.3	44.0	41.6	50.4	51.9	54.7	46.2	42.3	37.0	56.8	73.4
	Providence	36.6	40.3	53.1	58.4	59.7	59.4	61.1	61.4	51.9	44.2	67.5	83.5
	Washington	23.9	22.7	31.0	40.9	36.6	39.4	43.8	43.7	38.7	34.3	47.8	70.6
C)	The 8 counties of CONNECTICUT												
	Fairfield	30.9	24.5	43.7	47.1	56.3	50.7	48.9	43.3	38.9	29.8	46.6	60.8
	Hartford	35.9	31.4	46.4	49.9	61.2	56.5	57.3	54.3	49.4	41.9	58.9	73.0
	Litchfield	32.5	30.0	36.0	41.9	48.1	46.0	44.4	41.0	36.1	30.1	46.1	65.8
	Middlesex	33.1	29.9	39.7	46.3	52.9	49.2	48.6	47.5	41.5	35.2	40.1	67.5
	New Haven	36.5	34.4	50.5	52.4	60.5	55.0	53.1	50.2	45.1	37.0	58.0	69.1
	Tolland	31.0	30.3	39.6	40.0	52.8	50.5	48.5	46.9	41.2	36.5	48.6	69.0
	Windham	37.1	36.6	48.5	53.1	52.4	55.4	55.3	53.0	46.4	40.4	57.6	73.5
D)	The 3 counties of DELAWARE												
	Kent	52.6	50.1	40.7	57.2	56.5	53.3	52.8	49.0	41.3	47.5	50.1	59.0
	New Castle	39.9	33.5	32.0	45.2	55.6	55.7	56.8	50.3	48.2	44.2	51.4	62.6
	Sussex	46.4	46.1	35.7	51.3	51.6	52.7	47.3	45.8	46.6	45.2	48.6	55.7

exhibit 37 of chapter 17

Local mean time of moonrise, July 1971
(hours and minutes, 24-hour clock, meridian of Greenwich)

Date	Latitude						
	0°	10°	20°	30°	40°	50°	60°
1	12 21	12 30	12 38	12 48	13 01	13 18	13 45
2	13 04	13 16	13 29	13 43	14 02	14 28	15 11
3	13 49	14 05	14 21	14 41	15 05	15 40	16 41
4	14 38	14 51	15 17	15 40	16 10	16 53	18 14
5	15 31	15 52	16 15	16 41	17 14	18 04	19 44
6	16 28	16 50	17 13	17 41	18 15	19 07	20 55
7	17 27	17 48	18 10	18 36	19 10	19 58	21 33
8	18 25	18 44	19 04	19 27	19 56	20 37	21 51
9	19 21	19 36	19 53	20 11	20 34	21 06	21 59
10	20 15	20 26	20 37	20 50	21 06	21 29	22 03
11	21 05	21 12	21 18	21 26	21 35	21 47	22 05
12	21 55	21 56	21 57	21 59	22 01	22 03	22 07
13	22 43	22 39	22 36	22 31	22 26	22 19	22 08
14	23 33	23 24	23 15	23 05	22 52	22 36	22 10

SOURCE

The American Ephemeris and Nautical Almanac for the Year 1971, page 402.

exhibit 38 of chapter 17

Relative population of each age of Georges Bank Haddock
in terms of numbers caught per day

	Age in years							
Year	1	2	3	4	5	6	7	8
1931	147	691	158	439	699	466	256	132
1932	11	210	2,829	275	413	323	146	74
1933	44	986	720	1,145	249	193	145	67
1934	141	966	1,108	690	678	241	125	60
1935	202	1,704	1,306	574	509	428	97	74
1936	157	1,752	1,834	920	402	236	222	41
1937	140	1,233	1,327	698	535	251	119	65
1938	165	2,590	988	489	234	199	114	31
1939	95	1,774	2,416	640	271	123	108	42
1940	524	1,116	1,689	1,018	309	184	93	28
1941	144	3,298	1,275	1,046	752	233	123	40
1942	94	3,036	2,567	1,037	624	362	158	36
1943	11	1,026	3,470	1,551	530	492	149	61
1944	14	135	1,412	2,609	948	416	95	91
1945	25	1,663	420	1,244	1,218	485	194	61
1946	24	856	1,992	400	854	562	217	49
1947	18	1,996	1,189	863	250	314	180	89